الإحصـــــاء

وصف وتحليل وتطبيق

فــى

مجال السياحة والفنادق

الجـــزء الأول

دكتور / محمد الزنفلى

بسم الله الرحمن الرحيم

(وما توفيقى آلا بالله علية توكلت و الية أنبت)

صدق الله العظيم

الآية ١٨٧ سورة هود

الباب الأول

مقدمة عن علم الاحصاء

الباب الأول

مقدمه عن علم الإحصاء

١ – نشأة علم الإحصاء وتطوره :-

نشأ علم الإحصاء منذ زمن بعيد وكان مقتصرا على إيجاد (العدد) أو الحصر للموارد وأهما العنصر البشرى وذلك بهدف فرض الضرائب أو لخدمة الأغراض العسكرية ، ولذلك أطلق على علم الإحصاء بأنه علم العد أو علم الحساب السياسي . ولما كان القائمون على إيجاد هذا العدد هم موظفو الدولة فقد أطلق على علم الإحصاء اسم Statistics نسبة إلى الدولة State .

كما نشأ علم الإحصاء عند العربى القديم من إجراءه لعملية العد للأغنام بواسطة الحصى ، فقد كان الأعرابى أثناء خروج أغنامه من الحظيرة فى الصباح يقوم بوضع حصوه فى كيس معه مقابل كل رأس من الأغنام تخرج من الحظيرة ، وبهذا الأسلوب يكون قد صار فى الكيس عددا من الحصوات مساويا لعدد الأغنام التى خرجت من الحظيرة للرعى. وعندما يأتى المساء ويستقبل الأعرابى أغنامه للمبيت فى الحظيرة يقوم بإخراج حصوه من الكيس مقابل كل رأس من الأغنام ، وفى نهاية هذا العمل يتحدد لدى الأعرابى الأتى : –

١– إذا بقيت حصوة فى الكيس لم يقابلها رأس من الغنم خرج الاعرابى للبحث عن ضالته ووضع فى اعتباره التوقعات والفروض التالية : –

١- هل ضلت الطريق ؟
٢- هل هتك بها حيوان ؟
٣- هل سرقها لص ؟
٤- هل أصابها مرض و أقعدها فى المرعى ؟

ثم يحاول الإعرابي التحقق من صحة كل فرض حتى يقف على السبب الحقيقى لنقص عدد أغنامه ومن ثم يتمكن من وضع الإجراء المناسب لمنع هذا السبب مستقبلا .

٢- أما إذا انتهى الحصى الذى بالكيس ولا زالت توجد واحدة أو اكثر من الأغنام فان هذا يعني وجود زيادة فى العدد وهنا يبدأ الإعرابي فى وضع التخمينات والفروض التالية :-
١- هل أغنامه أنجبت واحدة ؟
٢- هل جذب قطيع أغنامه واحدة أو اكثر من غنم الجيرة ؟

ثم يحاول الإعرابي التحقق من صحة كل فرض حتى يقف على السبب الحقيقي للزيادة فى عدد الأغنام مما يمكنه من وضع الحل الصحيح لتلك الزيادة فى العدد . وهكذا كان الإحصاء عند العربى القديم يعنى العدد وذلك بهدف التنبؤ واتخاذ القرارات التخطيطية وبما يساعد على تفسير الفروض والتحقق من صدقها .

وقد ذكر الخالق العظيم دور الإحصاء فى التخطيط :-

١- ففي قصة " يوسف " أن قومه سيأتي عليهم سنوات عجاف يقل فيها المحصول ، وسنوات سمان يزيد فيها المحصول وهنا وجب العد والتخطيط للاحتياط لسنين القحط بادخار جزء من إنتاج سنين الوفاء.
٢- وقال الله سبحانه " **واعدوا لهم ما استطعتم من قوة ومن رباط الخيل** " وبالطبع هذا الإعداد يتطلب العد للموارد والإمكانيات ...

٣- كما علمنا القران الكريم من تعاقب الليل والنهار ... علم الحساب ، والإحصاء هو فن التعامل مع الرقم و الحساب .

وقد وردت كلمة الإحصاء فى القران الكريم :-

- " لقد أحصاهم وعدهم عدا "
- " و أحاط بما لديهم وأحصى كل شيء عددا "
- " وكل شيء أحصيناه فى إمام مبين "
- " ولا يغادر صغيرة ولا كبيرة إلا أحصاها "
- " و إن تعدوا نعمة الله لا تحصوها "
- " فطلقوهن لعدتهن وأحصوا العدة "

و أيا كان المنشأ غربى أم عربى فقد تطور فكر الإنسان عن مفهوم الإحصاء كما يلى : -

١. إن الإحصاء هى جمع وعرض وتلخيص البيانات Reduction of data وقد ظل هذا المفهوم ما يقرب من نصف قرن .

٢. إن الإحصاء هى الدراسة الرقمية للمجتمعات Populations والمجتمع الإحصائي هو مجموعة من المفردات (أفراد ، أشياء ، أدوات ، . . . .) تجمع بينهما صفة أو صفات مشتركة بشرط عدم التشابه أو التجانس الكامل للمفردات فى الصفة أو الصفات محل الدراسة بل يجب أن تتسم بالتغير .

فإذا كان المجتمع الإحصائي تتسم مفرداته بالتشابه أو التجانس الكامل لمفرداته فلا داعي للدراسة الإحصائية لهذا المجتمع من خلال كل مفرداته أو حتى من خلال عينة مسحوبة منه بل من خلال مفردة واحدة . وتعتبر المجتمعات الإحصائية والتى تتسم مفرداتها بالتغير فى الصفة أو

الصفات المعينة هى الممثلة للأغلبية الساحقة لظواهر ومشكلات اوجه الحياة .

وجدير بالذكر أن الجداول الرياضية مثل جدول الفائدة المركبة أو جدول اللوغاريتمات وغيرها من الجداول الرياضية لا تعد مجتمعات إحصائية رغم كونها مجتمعات رقمية ، والسبب أن مثل هذه الجداول الرياضية تتحدد خصائصها بواسطة القانون الرياضي .

ومن أمثلة المجتمعات الإحصائية المجتمع الإحصائي عن عدد المواليد أو الوفيات، الصادرات ، الدخل السياحي ، عدد الليالي السياحي ، عدد الفنادق ، ... .

٣. أن الدراسة الإحصائية ليست بالضرورة هى الدراسة الرقمية للمجتمع الإحصائي بأكمله وإنما باستخدام عينة Sample مسحوبة من هذا المجتمع تحت شروط معينة ، وتسمى تلك الدراسة الإحصائية بالاستدلال الإحصائي Statistical inference اى معرفة خصائص المجتمع الكبير من خلال دراسة إحصائية لعينة مسحوبة منه ، هذه المعرفة يتم قبولها أو رفضها بدرجة ثقة معينة ، وتقاس هذه الثقة أو الشك (عدم التأكد) Measure of uncertainty بواسطة الاحتمالات .

فعند تقدير Estimat معلمة المجتمع Parameter باستخدام إحصاء العينة Statistic المسحوبة منه ، نفرض أننا بصدد تقدير نسبة الطالبات فى إحدى الكليات ، وقمنا بسحب عينه عشوائية* حجمها ١٠٠٠

*يقصد بالعشوائية اختيار مفردات العينة من مفردات المجتمع بأسلوب يسمح بإعطاء كل مفردات المجتمع فرصة معلومة للظهور فى العينة .

من مجموع طلاب الكلية وتبين أن بالعينة ٤٠٠ طالبة ، فإن نسبة الطالبات فى العينة تساوى ٠,٤٠ وتسمى هذه النسبة بإحصاء العينة وتستخدم فى تقدير نظيرها اى نسبة الطالبات فى مجتمع الكلية اى المعلمة، مع اقتناعنا ان هذا التقدير يحمل نسبة خطأ مقبولة ، فنقول مثلا ان الفرق بين (إحصاء العينة) و (معلمة المجتمع) لن يتعدى عن ٠,٠٢٧ بدرجة ثقة ٠,٩٥ ، وسنرى فيما بعد كيفية إجراء مثل هذه الحسابات وذلك عند دراسة نظرية توزيع المعاينة ** التى تفيدنا فى تقييم هذا الخطأ ومقداره .

كذلك عند اختبار فرض Test of Hypothese يتعلق بمعلمة المجتمع ، نفرض أننا بصدد اختبار مصل جديد يقى الأطفال من موض الشلل فقمنا بسحب عينة حجمها ٤٠٠٠٠ طفل وتقسيمها إلى مجموعتين متساويتين على أن يتم تطعيم المجموعة الأولى بالمصل وترك المجموعة الثانية بدون تطعيم ، وبتسجيل حالات الإصابة بمرض شلل الأطفال فى المجموعتين يتبين أنها ٣٥ حالة فى المجموعة الأولى ، ١٤٢ حالة فى المجموعة الثانية ، أى يوجد فرق بين المجموعتين وبناءا على ذلك تقرر تعميم التطعيم بهذا المصل مع إقتناعنا بأن هذا القرار يحمل نسبة خطأ مقبولة وسنرى فيما بعد مثل هذه الاختبارات الإحصائية .

٤. إن الدراسة الإحصائية لا تقتصر على وصف نمط الاختلاف لظاهرة واحدة فقط ، بل تتعدى ذلك إلى دراسة العلاقة بين ظاهرتين أو اكثر بهدف تفسير الاختلاف بين تلك الظواهر أو التنبؤ بسلوكها فى المستقبل ، ويتحقق ذلك من خلال الارتباط و الانحدار.

** Theory of sampling distribution .

مما سبق يعرف علم الإحصاء بأنه مجموعة من طرق علمية تختص بالحصول على البيانات عن الظواهر المختلفة (طبيعية أو اجتماعية أو سياسية) ومراجعتها وتبويبها وتهذيبها وتحليلها لتفسيرها واستنباط الحقائق المتعلقة بها ومعرفة القوانين التى تسير عليها .

وجدير بالذكر أن علم الإحصاء يدرس الظواهر من خلال عدد كبير من المشاهدات حتى يمكن تطبيق قانون الأعداد الكبيرة التى تعتبر حجر الزاوية فى القوانين الإحصائية وبدونه لا يمكن استخدام النتائج المتحصل عليها فى التعميم على الظاهرة محل الدراسة .

كما تجدر الإشارة إلى أن لكل ظاهرة جانبان ، جانب كمي و آخر كيفي أو نوعي ، ويهتم علي الإحصاء بدراسة الجوانب الكيفية أو النوعية للظواهر وذلك باستخدام أساليب كمية ، وهذا ما يميز الإحصاء عن الرياضة التى تهتم بالكم فقط بغض النظر عن الكيف .

ويعتبر الإحصاء الرياضى (الإحصاء البحت أو النظرية الإحصائية) أساسا هاما فى الإحصاء التطبيقى حيث يهتم الإحصاء الرياضى بوضع القوانين والمبادئ التى تستخدم فى الإحصاء التطبيقية ، كما أن الإحصاء التطبيقى يضع أمام الإحصاء الرياضى المشكلات العلمية التى تتطلب حلولا بل والمزيد من الطرق الإحصائية المتطورة لحلها ، وعلى ذلك فكلاهما يثرى الآخر ، فنظرية الاحتمالات تمد النظرية الإحصائية بالعديد من المبادئ والقوانين التى توصل إلى نتائج وقرارات معينة . "

أهمية علم الإحصاء :

١.في التخطيط :

يحتاج المخطط إلى بيانات و أرقام إحصائية تساعده على كشف الواقع أو المستقبل وذلك بتلمس الاحتياجات والتعرف على الإمكانيات ، ومن ثم يتمكن المخطط من وسم خطته ، بل أن المخطط بعد تنفيذ الخطة يكون في اشد الحاجة إلى بيانات و أرقام إحصائية تمكنه من إجراء التقييم Evolution للنتائج التي توصل إليها وبالتالي التعرف على مدى تحقيق الهدف المنشود .

٢.في البحث العلمي :

البحث العلمي بصفة عامه يعني التفكير المنظم في سلوك الظاهرة محل الدراسة ومحاولة إيجاد تفسير مقبول لهذا السلوك حتى يمكن السيطرة عليه وتوجيهه الوجه المقصودة تجاه تلك الظاهرة .ويجرى هذا التفسير من خلال وضع الفروض النظرية – (والفروض النظرية هي النظرية العلمية في مجال الظاهرة) – ثم تجميع البيانات والإحصاءات عن تلك .الفروض ثم التحليل الإحصائي والتوصل إلى قرار بشأن تلك الفروض من حيث قبولها أو رفضها . فإذا كانت الظاهرة (المشكلة) محل الدراسة هي دراسة دالة الطلب مثلا لإحدى السلع في سوق ما ، فان النظرية الفرضية (أي النظرية الاقتصادية في هذه الحالة) توضح أن المستهلك الرشيد هو الذي يسعى للحصول على أقصى إشباع ممكن في ظل دخلة المتاح ، وبالتالي فان الطلب على سلعة ما يتأثر بالعوامل التالية :

- التأثير السلبى لسعر تلك السلعة .
- التأثير الإيجابي لدخل المستهلك لهذه السلعة .
- التأثير الإيجابي لأسعار السلع البديلة .
- التأثير الإيجابي لأذواق المستهلكين .

إلى غير ذلك من العوامل المؤثرة ، وبعد ما يتم جمع البيانات الإحصائية عن تلك العوامل المؤثرة ثم إجراء التحليل الإحصائي والاقتصادي يتم قبول أو رفض تأثير تلك العوامل على المشكلة محل الدراسة وهى دالة الطلب للسلعة المعنية .

٣.في بحوث العمليات :

حدث عام ١٩٣٨ أن بريطانيا استحدثت { أسلوب علمى } يهدف إلى تحليل خسائرها العسكرية فى الحرب العالمية الثانية بغرض تخفيض حجم خسائرها العسكرية ، وقد اعتمدت فى ذلك على خبرة وجهود العلماء والمتخصصين فى مجالات : -

- الرياضة .
- الإحصاء .
- العلوم الطبيعية .

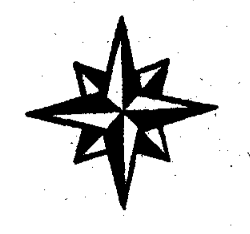
وقد أطلق على هذا الأسلوب العلمى اسم " بحوث العمليات " Operations Research ، ثم شاع استخدام هذا الأسلوب فى مجالات متعددة مدنية منها :

- البرمجة الخطية Linear programming
- التحليل الشبكى Net work analysis
- نظرية المباريات Theory of games

- أسلوب المحاكاة Simulation model
- نظم التخزين Inventory systems

وقد ساعد تطور واستخدام الحاسبات الإلكترونية على تطور هذه الأساليب وانتشارها حتى اصبح كل منها فرع علمي مستقل بذاته .

ولقد ظهرت أهمية الإحصاء فى جميع دول العالم ، و أصبحت كل دولة تصدر التشريعات التى تنظم النشاط الإحصائي فيها ، ففى ج.م.ع يعتبر الجهاز المركزى للتعبئة العامة والإحصاء هو المسئول عن النشاط الإحصائي بصفة خاصة بالإضافة إلى ما تقوم به الوزارات المختلفة من أنشطة إحصائية خاصة بها . وتعتبر وزارة السياحة بأجهزتها المختلفة ، والهيئة المصرية العامة لتنشيط السياحة ، والهيئة المصرية العامة للتنمية السياحية ، ووزارة الداخلية ممثلة فى مصلحة وثائق السفر والهجرة والجنسية ، ووزارة المالية هى الأجهزة التى تهتم بالإحصاءات السياحية فى مصر .



الباب الثانى

جمع البيانات الاحصائية

جمع البيانات الإحصائية

يشتمل هذا الباب على الفصول التالية : -

الفصـــل الأول : مفاهيـــم أساسيــــه

الفصل الثاني : مصادر البيانات وطرق جمعها

- المصادر الثانويــــــة
- المصادر الأولية (الميدانية)

الفصل الثالث : الحصر الشامل والعينة فى جمع البيانات

الفصل الرابع : الأخطاء الشائعة فى جمــع البيانـــــات



الفصل الأول

مفاهيم أساسية

الطريقة العلمية فى البحث العلمى :-

تبنى الطريقة العلمية فى البحث على عدة عناصر أهمها :

١- مشاهدة الظاهرة **Observation**

٢-وضع الفروض المنطقية لأسباب الظاهرة **Logical Hypothesis**

٣-التوقع (الاستنباط أو التنبؤ) **Predication**

٤-اختبار الفروض **Testing of Hypothesis**

وتعتبر الطريقة العلمية هى الأكثر موضوعية Objective من غيرها كالتى تعتمد على التخمين Presumption ، فنتائج الطريقة العلمية لا تتأثر بشخصية الباحث أو بمشاعره أو ببيئته .

وتعتمد الطريقة العلمية على الأسلوب الرقمى (لغة الأرقام) للبيانات التى يتم تجميعها عن الظاهرة محل الدراسة أو البحث . فالأرقام التى تشير إلى ازدياد مبيعات مطعم سياحى معين تعبر عن ظاهرة ، والأرقام التى تشير إلى ارتفاع التكاليف التسويقية لخدمة سياحية معينة تعبر عن ظاهرة ، والأرقام التى تشير إلى تغير الأسعار........ .



المفردات الإحصائية :-

المفردات الإحصائية ليست بالضرورة أشخاص أو كائنات حية ، بل قد تكون سلع أو مصانع أو مزارع أو فنادق أو منازل أو انهار أو بحار أو طرق أو ...

ويشترط فى مجموعة المفردات الإحصائية محل الدراسة ان تكون ذات تجانس كامل . فإذا كنا بصدد طلبة قسم الفنادق فان جميع طلبة هذا القسم هم مجموعة مفردات إحصائية ، وإذا كنا بصدد طلبة السنة الثانية بقسم الفنادق فان طلبة السنة الثانية دون غيرهم من طلبة القسم مجموعة مفردات إحصائية .

الصفة الإحصائية :-

هى الخاصية التى يراد دراستها لمجموعة المفردات الإحصائية، فمثلا إذا كنا بصدد مجموعة المفردات الإحصائية لطلبة قسم الفنادق ، و أردنا دراسة صفة النوع (بنت أم ولد) لهؤلاء الطلاب ، فان صفة النوع تسمى صفة إحصائية . وقد تكون الصفة الإحصائية هى الطول أو الوزن أو درجات النجاح فى مادة الإحصاء مثلا ...

ويشترط فى الصفة الإحصائية عدم التشابه أو التجانس التام بل تتغير من مفردة لأخرى داخل مجموعة المفردات الإحصائية محل الدراسة : فمثلا إذا كانت مجموعة المفردات الإحصائية هى طلاب جامعة الأزهر و أريد دراسة صفة الإسلام عند هؤلاء الطلاب ، فهذه ليست صفة إحصائية إذ أن جميع هؤلاء الطلاب مسلمون ، أي أن هذه الصفة ذات تجانس تام . لكن صفة درجة النجاح فى مادة الإحصاء لطلبة السنة الثانية بقسم السياحة تعتبر صفة إحصائية إذ أن درجة النجاح تتغير من طالب لأخر .



المتغير الإحصائي :

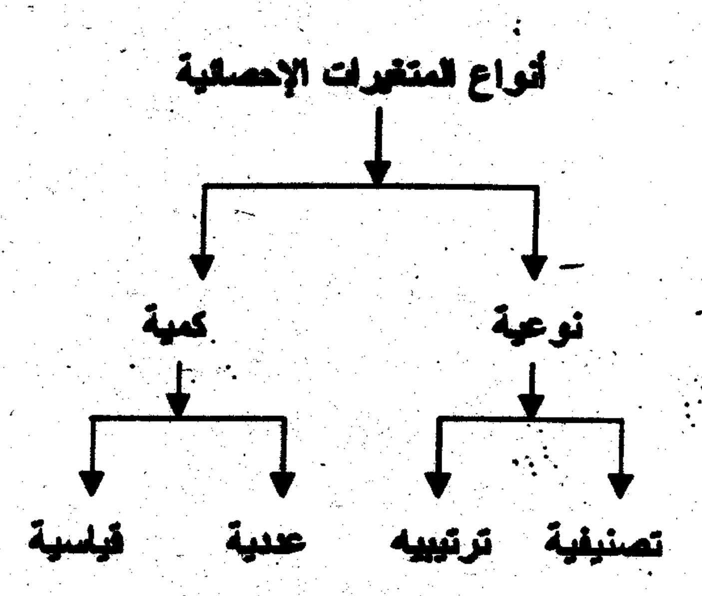
يعبر المتغير الإحصائي Variable عن الصفة الإحصائية التى يراد دراستها لمجموعة المفردات الإحصائية . وقد تهتم الدراسة الإحصائية بصفة واحدة لمجموعة المفردات الإحصائية، وهنا تكون البيانات خاصة بمتغير إحصائي واحد ، وقد تهتم الدراسة الإحصائية بصفتين أو اكثر لمجوعة المفردات ، وهنا تكون الدراسة الإحصائية لعدة متغيرات .

والمثال التالى يوضح عدة متغيرات لمجموعة مكونة من ١٠ طلاب بقسم الإرشاد السياحى :

رقم الطالب	النوع	الجنسية	الوزن بالكيلو جرام	الطول بالسنتيمتر	المسافة التى يقطعها للجامعة
١	ولد	مصري	٨٠	١٧٥	٢٧
٢	ولد	فلسطيني	٧٥	١٧٠	٧
٣	بنت	فلسطينية	٧٠	١٣٠	٥
٤	ولد	مصري	٩٠	١٤٠	٣٥
٥	بنت	مصرية	٧٥	١٥٠	٢٥
٦	بنت	مصرية	٦٥	١٣٠	٢٠
٧	بنت	سودانية	٦٠	١٥٥	١٧
٨	ولد	مصري	٧٨	١٩٠	٣
٩	ولد	مصري	٨١	١٦٠	٧٥
١٠	ولد	تونسي	٧٦	١٦٠	١٠

وتنقسم المتغيرات الإحصائية إلى عدة أنواع يمكن عرضها فى الرسم التخطيطي التالي : -





والمتغيرات النوعية Categorical or qualitative variables هـى التى يتم فيها التفرقة بين صور المتغير الإحصائي على أساس نوعى ، مثال ذلك صفة الحالة الاجتماعية وصورهـا (أعـزب ، متزوج ، مطلق، أرمل) لسكان ج.م.ع ، وأيضا صفة تقديرات النجاح وصورها (ممتاز ، جيد جدا ، جيد ، مقبول ، ضعيف ، ضعيف جدا) لطلبة كلية ما . وإذا كان الاهتمام بالمتغير النوعى ينصب فقط على تصنيف البيانات كما فى المثال الأول سمى متغير تصنيفى Nominal ، أما إذا كان الاهتمام بالتصنيف والترتيب سمى متغير ترتيبي Ordinal كما فى المثال الثاني فهذه الصور التصنيفية لدرجة النجاح يمكن ترتيبها تصاعديا أو تنازليا ، وتجدر الإشارة إلى انه قد تستخدم الأرقام فى حالة المتغيرات النوعية لكن بأهداف مختلفة ، فمثلا فى حالة البيانات النوعية التصنيفية كنوع الشخص بنت أم ولد يتم إعطاء الرقم (١) لكون النوع بنت ، و إعطاء الرقم صفر لكون النوع ولد وذلك بهدف تسهيل التعامل مع البيانات ليس إلا ، وفى حالة البيانات النوعية



الترتيبية كالوضع الاجتماعي للأسرة و إعطاء الرقم (١) للطبقة العليا والرقم (٢) للطبقة المتوسطة والرقم (٣) للطبقة الدنيا وذلك بهدف تسهيل التعامل مع البيانات بالإضافة إلى إظهار الوزن النسبي لكل صور المتغير .

أما المتغيرات الكميه Numerical or Quantitative Variables فهى التى يتم فيها التفرقة بين صور المتغير الإحصائي على أساس كمى . ويسمى المتغير الكمى متغير عددى إذا كان العد Counting هو المستخدم فى دراسة المتغير ، ويعرف المتغير فى هذه الحالة بأنه متغير متقطع Discrete، مثال ذلك عدد أفراد الأسرة، عدد السائحين إلى مصر ، عدد الغرف فى السكن ، وعدد المقررات التى يدرسها الطالب ، أي أن المتغير المتقطع هو الذى يأخذ قيما من بين مجموعة الأعداد الصحيحة {صفر، ١ ، ٢، ٣ ، ٤ ، .... } كما يسمى المتغير الكمى بمتغير قياسي إذا كان القياس Measurement هو المستخدم فى دراسة المتغير ، ويعرف المتغير فى هذه الحالة بالمتغير المتصل Continuous، مثال ذلك قياس الوزن للشخص أو قياس الطول أو قياس دخل الأسرة ، فكل هذه القياسات لا تتحقق غالبا فى رقم صحيح و إنما برقم صحيح وكسر الواحد الصحيح.

المجتمع الإحصائي المحدود (Finite) :-

هو المجتمع الذى يمكن معرفة عدد مفرداته مثل عدد الفنادق فى مصر ، عدد السائحين الوافدين إلى مصر ... وبالتالى يمكن دراسة أي صفة سواء بالحصر الشامل لجميع مفرداته أو بعينه مفردات مسحوبة منه .



المجتمع الإحصائي غير المحدود (Infinite) : -

هو المجتمع الذى لا يمكن معرفة عدد مفرداته مثل السمك فـى البحر الأحمر وبالتالى لا يمكن دراسة اى صفة لمفردات هذا المجتمـع الإحصائي إلا من خلال عينه مفردات مسحوبة منه .



الفصل الثانى

مصــادر البيانات الإحصــائية
وطـــــرق جمعهـــــا

إن تحديد المشكلة لظاهرة ما يؤدى إلى تحديد نـــوع البيانـــات المطلوب جمعها ، وبالتالى تتحدد المصادر التى يمكن الحصول منـــها على تلك البيانات . ومصـادر البيانات الإحصائية قد تكـــون مصـــادر ثانوية أو مصادر أولية .

المصادر الإحصائية الثانوية : Secondary Sources

الإحصاءات الثانوية هى بيانات تاريخية أو ما يطلـــق عليـــها بيانات السلسلة الزمنية أي بيانات لعينة طويلة من الزمن لظاهرة مـــا . ومصادر هذه البيانات هى السجلات التى تحتفظ بها المنشأة (فنـــدق ، شركة سياحية ، مصنع ، متجر ، مزرعـــة ، ....) أو تحتفـــظ بـــها الجهات الإحصائية المتخصصة فى الدولة .

وللمصادر الثانوية عدة مزايا أهمها توفـــير الوقـــت والجـــهد للباحث ، لا تتطلب تكلفة كبيره للحصول عليها ، إلا انه يعاب عليـــها بأنها قد تكون غير متوفرة ، أو قد تكون متوفرة لكن درجة الثقة فيـــها غير كافية . ولمواجهة هذه العيوب يضطر الباحث إلى جمع البيانـــات بنفسه من الميدان وهذا ما يعرف بالبيانات الأولية .



المصادر الإحصائية الأولية (الميدانية): Primary Sources

ترجع أهمية هذه البيانات إلى أنها غير متاحة فى سجلات المنشأة أو حتى خارجها ، كما أنها قد تتعلق بمشكلة خاصة لم يسبق للمنشأة أن تعرضت لها وهنا يقوم الباحث بجمع هذه البيانات بنفسه من المفردات محل الدراسة ، ويتم جمع هذه البيانات بثلاث طرق : -

١- طريقة المشاهدة .

٢- طريقة الاستقصاء .

٣- طريق التجريب.

أولا : طريقة المشاهدة Observation

الملاحظة أو المشاهدة هى الطريقة العلمية الرئيسية التى يمكن أن يتبعها الباحث للحصول على معلومات عن الظاهرة محل الدراسة أو البحث . والمشاهدة تكون علميه إذا كانت تخدم هدفا علميا محددا ، وان يتم تسجيلها بطريقة منتظمة ، وألا تكون لمجرد الفضول الطبيعى للإنسان ، وان تكون موضوعية يمكن الاعتماد عليها والتحقق منها . كما أن المشاهدة كطريقة في جمع البيانات يقوم عليها الباحث إذا رأى أن المفردة موضوع البحث قد تغير سلوكها إذا ما شعرت أنها موضع دراسة وبحث فهنا يقوم الباحث بالمشاهدة وجمع المعلومات دون إشعار الشخص موضع الدراسة بأنه تحت الملاحظة . وتتحقق المشاهدة بطريقتين هما طريقة المشاهدة الشخصية وطريقة المشاهدة بالآلة .

* **طريقة المشاهدة الشخصية : -**

فيها يقوم الباحث أو جامع البيانات بملاحظة الشخص موضع



الدراسة وتدوين الملاحظات ثم ترجمتها إلى بيانات ومعلومات .

* **طريقة المشاهدة بالآلة : -**

فيها يتم تركيب آلات تصوير أو كاميرات تليفزيونية فى أماكن الشراء لتصوير حركة العملاء و إبراز سلوكهم وكيفية اختيارهم للسلع المختلفة ثم تدوين ذلك فى شكل معلومات وبيانات .

مزايا المشاهدة كطريقة فى جمع البيانات : -

- فيها يتم تسجيل التصرف أو السلوك عند وقوعه عن المفردة .
- تستخدم فى الحالات التى يصعب الحصول فيها على بيانات مباشرة من الشخص موضع الدراسة .
- تستخدم عادة فى الدراسات الأولية الاستكشافية .
- تستخدم فى الحصول على بيانات ومعلومات إضافية تساعد فى تفسير الظواهر .

<u>وكمثال تطبيقى</u> : - يتم استخدام المشاهدة فى جمع البيانات عندما يراد معرفة عدد الأفراد الذين يترددون على سوق معين أو محل تجارى معين ، أو نسبة الذين اشتروا إلى الذين دخلوا هذا المكان .

عيوبها : -

- اكثر تكلفة من الطرق الأخرى .
- ما زال استخدامها فى دراسة الأسواق اقل من الطرق الأخرى .

ثانيا: طريقة الاستقصاء Questionnaire

يطلق على طريقة الاستقصاء فى جمع البيانات طريقة الاستبيان أو استطلاع الرأي ، وفى هذه الطريقة يتم توفير استمارة استبيان لكل مفرده من مفردات المجتمع الإحصائي محل الدراسة لتقوم كل مفرده بالإجابة على الأسئلة التى تتضمنها الاستمارة ،وقد يقوم جامع البيانات شخصيا بتدوين إجابات المستقصى منه فى استمارة الاستبيان وذلك بالمقابلة الشخصية مع المستقصى منه ، أو قد يقوم المستقصى منه بنفسه بتدوين الإجابات فى استمارة الاستبيان التى تصل إليه بالبريد .

وتعتبر طريقة الاستقصاء أكثر طرق جمع البيانات شيوعا وذلك للأسباب التالية :-

- الاستقصاء طريقة مرنة يمكنها أن تغطى أي مجال .
- طريقة الاستقصاء أسرع وأقل تكلفة من الطرق الأخرى فى جمع البيانات ، ويعاب على طريقة الاستقصاء بـ :

- عدم رغبة المفردات فى التعاون مع القائمين بالدراسة حيث يمتنع المستقصى منه عن الإجابة على الأسئلة المطلوبة .
- قد لا يكون فى مقدور المستقصى منه الإجابة على الأسئلة الموجهة اليه لسبب أو لأخر مثل النسيان أو لعدم معرفته الإجابة الصحيحة وعدم ملائمة طريقة الاستقصاء فى جمع البيانات عن دوافع السلوك ، ويمكن التغلب على هذه العيوب باتباع الطريقة المناسبة من طرق الاستقصاء وهى طريقة المقابلة



الشخصية مع المستقصى منه أو طريقـــة إرســـال استمارة الاستبيان إلى المستقصى منه .

* طريقة المقابلة الشخصية **Personal interview**

فى هذه الطريقة يقوم جامع البيانات بتوجيه الأســـئلة التـــى تشتمل عليها استمارة الاستبيان إلى كـــل مفـــرده مـــن مفـــردات المجتمع الإحصائي محل الدراسة ثم يقوم بتسجيل إجابات الأفـــراد فى الأماكن المخصصة بالاستمارة . فمثلا عنــــد إجــــراء تعـــداد السكان يقوم العداد (الشخص جامع البيانات) بمقابلة رب كـــل أسرة للحصول منه على البيانات والمعلومــــات المطلوبـــة عـــن أسرته، وفى بحوث التسويق يقوم جامع البيانات بمقابلة المستهلكين للحصول منهم علي بيانات تتعلق بآرائهم حول خصائص الســـلع المختلفة ، وفى دراسة عن أسباب رسوب الطلبة فى الامتحانـــات

- يمكن الحصول على قدر اكبر من البيانات
- تمكن من تلافى التضارب فى الإجابة على اكثر من سؤال .
- تمكن من تكوين ملاحظات شخصية
- تكتسب هذه الطريقة أهمية خاصة فى المواقف التى يرتفع فيها مستوى الأمية أو يقل فيها مستوى الوعي الإحصائي بين مفردات المجتمع الإحصائي محل الدراسة .

ويعاب عليها :-

- وجود تحيز فى البيانات ناشئ عن عدم خبرة أو عدم أمانة جامعي البيانات ، فقد يهمل جامعو البيانات اتباع التعليمات المعطاة لهم أو قد يتعاملون مع مفردات الدراسة بأسلوب يؤثر فى نوعية إجاباتهم أو قد يخطئون عند تسجيل البيانات المعطاة لهم .
- قد تكون هذه الطريقة غير مناسبة إذا كان المستقصى منه يريد وقت للتفكير فى السؤال قبل الإجابة عليه ، أو إذا كانت الأسئلة فيها ما يحرجه .

- **طريقة إرسال استمارة الاستبيان للمستقصى منه**

أي طريقة الاستبيان الذاتى Self-Enumeration

فى هذه الطريقة يتم توفير استمارة استبيان لكل مفرده من مفردات الدراسة مع إضافة تعليمات واضحة عن كيفية قيام المستقصى منه بالإجابة على أسئلة استمارة الاستبيان المرسلة إلية



فمثلا فى دراسة عن خريجى السياحة والفنادق تم إرسال استمارة استبيان لكل خريج وذلك بهدف الحصول على إجابات عن وظائفهم وأنشطتهم المختلفة منذ التخرج ، وفى دراسة تقوم بها شركات الأدوية بإرسال استمارة استبيان للأطباء وذلك للحصول على بيانات تتعلق بآرائهم وملاحظاتهم حول الأدوية التى تنتجها هذه الشركات .

وتتميز طريقة الاستبيان الذاتى فى جميع البيانات بالآتي : -

- تمكن من تفادى أخطاء التحيز التى قد تنشأ من تدخل جامعى البيانات .
- تصلح فى حالة احتواء استمارة الاستبيان على أسئلة بها حرج للمستقصى منه إذا ما وجهها إليه جامع البيانات شخصيا .

ويعاب عليها بـ :

- لا يمكن إستخدام هذه الطريقة إذا كان مستوى الأمية مرتفعا بين مفردات المجتمع محل الدراسة .
- تتسم هذه الطريقة بانخفاض معدلات استجابة المستقصى منهم فى ملا استمارة الاستبيان و إعادتها وذلك لإهمال المستقصى منه فى الإجابة على أسئلة الاستمارة أو لانشغاله أو لعدم وجود حافز له ، لذلك تحرص هيئات جمع البيانات على الاتصال المتتالى بالمستقصى منهم وحثهم على الاستجابة . وتجدر الإشارة إلى أن انخفاض معدلات الاستجابة هذه يترتب عليها أن يكون عدد المستقصى منهم قليل ومن ثم يصبحون عينة غير ممثلة



للمجتمع الإحصائي محل الدراسة مما يوجد تحيز في البيانات التى تجمع ، ومن الواضح أن هذا التحيز يختلف فى نشأته عن التحيز السابق والناشئ من تدخل جامعي البيانات .

- لا تمكن من الحصول على قدر كبير من البيانات حيث مقدار البيانات الناتج منها اقل .
- اقل مرونة .
- يتعذر معها استيضاح سؤال معين يراه المستقصى منه .

إعداد استمارة الاستبيان

Questionnaire Design

يعتمد نجاح عملية جمع البيانات بشكل أساسى على جودة استمارة الاستبيان ، وفيما يلى بعض الاعتبارات التى يجب مراعاتها عند إعداد استمارة الاستبيان :

١. أن يهتم الباحث بالتسلسل المنطقى فى ترتيب الأسئلة وذلك بان يبدأ بمجموعة الأسئلة الخاصة بتعريف المفردة محل البحث مثل السن والحالة الاجتماعية ، ثم الأسئلة المتعلقة بهدف البحث على أن تؤجل الأسئلة الصعبة والمثيرة للجدل حتى النهاية .

٢. مراعاة اختلاف نوعية الأسئلة باختلاف الهدف من عملية الاستقصاء :

- فإذا كان الهدف من الاستقصاء هو جمع بيانات عن حقائق مثل عدد أفراد الأسرة ومؤهلاتهم فان الأسئلة تكون مفتوحة أو مغلقة .

- وإذا كان الهدف من الاستقصاء هو جمع بيانات عن أراء مثل ما رأيك في مزايا نوع معين من الصابون فيفضل أن تكون الأسئلة مغلقة حتى يمكن تنميط الإجابة قدر الإمكان ففي هذا المثال قد تكون الإجابة أحد الحالات التالية :

رخيص الثمن ☐ له رغوة كبيرة ☐ له رائحة جميلة ☐

- أما إذا كان الهدف من الاستقصاء هو جمع بيانات عن دوافع فان هذا الاستقصاء يعرف باستقصاء المدلولات ، وهذا النوع من الاستقصاءات يحتاج من المستقصى منه أن يتذكر ، فمثلا سؤاله كم عدد مرات استعمالك للسلع (س) فى الأسبوع ؟ أو لماذا تشترى ماركة معينه من السلعة (ص) ؟ أو لماذا أنت تدخن ؟ أو ما الدوافع وراء شرائك السلعة (ل) ؟ وهنا نلاحظ تدخل النواحى السلوكية وتأثيرها على الإجابة لذلك يجب توخى الحذر والدقة فى اختيار أسئلة هذا النوع من الاستقصاء منعا للتحيز المتعمد أو غير المتعمد من الشخص المستقصى منه .

وجدير بالذكر أن الأسئلة المفتوحة هى التى لا يتم فيها تحديد إجابات للمستقصى منه ، أما الأسئلة المغلقة هى التى يتم فيها تحديد إجابات بديلة للمستقصى منه كما فى مثال الصابون ، كما توجد أسئلة



مغلقة تكون الإجابة عليها بنعم أو لا وهى الأكثر استخداما فى بحوث التسويق .

٣. يجب التعريف الدقيق للمفاهيم التى تتضمنها استمارة الاستبيان فعند السؤال عن الأجر يجب تحديد ما إذا كان ذلك الأجر هو الأجر اليومى أم الأجر الاسبوعى أم الأجر السنوى .

٤. يجب الابتعاد عن الأسئلة التطفلية والتى تثير التحيز الشخصى وتكون الإجابة عليها معروفة مقدما .

٥. أن يشار فى استمارة الاستبيان أن البيانات لا تستخدم إلا فى أغراض البحث العلمى حتى يطمئن المستقصى منه ويميل إلى الاستجابة للإجابة على أسئلة استمارة الاستبيان.

٦. أن تصاغ الأسئلة بأسلوب سهل مع الابتعاد قدر الإمكان عن المصطلحات الفنية وان تستخدم اللغة العامية كلما أمكن ذلك

٧. أن يكون الأشخاص القائمين على ملأ استمارة الاستبيان سواء بالمقابلة الشخصية أو بالبريد على مستوى عالى من المهارة فى جمع البيانات لأنه قد تكون هناك ملاحظات ضمنية وليست صريحة وتؤثر على إجابة المستقصى منه .

ثالثا : طريقة التجريب Experimental Studies

تعتبر طريقة التجريب فى جمع البيانات اكثر ملائمة لدراسة العلاقات السببية بين المتغيرات وإيجاد البرهان على وجود هذه العلاقات ، وتطلب هذه الطريقة استخدام تجربة يتم تصميمها بأسلوب يسمح بالتحكم فى العوامل المختلفة المؤثرة على الظاهرة وأيضا قياس



تأثير هذه العوامل ، ويستخدم الإحصائيون مبدأ العشوائية في تصميم التجارب بحيث تكون التجربة غير متحيزة وان يتم فيها تمثيل جميع العوامل المؤثرة بشكل متوازن ، فمثلا عند دراسة تأثير عقار معين على إنقاص الوزن فانه يتم استخدام عينة من ١٢ فأر ثم يتم اختيار ٦ من بينها عشوائيا لتعاطى هذا العقار بينما يستمر الباقون فى نظامهم الغذائى العادى الذى يخلو من هذا العقار وذلك طوال الفترة التى يحددها الباحث ثم يقوم الباحث بجمع البيانات عن المشاهدات التى يلاحظها على المجموعتين من الفئران ثم يجرى التحليل الإحصائي اللازم للبيانات التى تم جمعها وذلك لإقرار مدى تأثير العقار من عدمه

وعلى ما سبق يلاحظ أن طريقة التجريب فى جمع البيانات ليس من السهولة تطبيقها فى العلوم الاجتماعية والاقتصادية حيث أن العوامل المؤثرة فى الظواهر الاجتماعية والاقتصادية يصعب التحكم فيها والسيطرة عليها كأذواق المستهلكين أو السلوك الإنساني أو العلاقات البشرية كما أنها لا تخضع لقواعد وقوانين ثابتة كما هو الحال فى العلوم الطبيعية .

وقد تستخدم طريقة التجريب فى العلوم الاجتماعية – وان كان بشكل نادر – فى بعض حالات دراسة الأسواق وذلك دون إشعار المفردات تحت البحث بأنها تحت التجربة والملاحظة ، فمثلا يمكن بحث اثر الإعلان على المبيعات بالإعلان عن السلعة فى أحد الأسواق ثم ترك أسواق أخرى دون إعلان ، ولو أن التحكم الإحصائي هنا غير كافي حيث قد تتدخل عوامل أخرى غير الإعلان فى التأثير على مبيعات السلعة فى كلا السوقين .

وقبل الانتهاء من تناول مصادر البيانات الإحصائية وطرق



جمعها فان الأمر يتطلب التعرض للإجراءات المتبعة عند جمع البيانات الإحصائية السياحية .

أولا : تسجيل البيانات الإحصائية السياحية عند القدوم : -

يتحقق ذلك فى الدفاتر الممسوكة (يدويا أو بالكومبيوتر) على الحدود والمنافذ الشرعية للدولة كالموانى البحرية والجوية والمنافذ البرية على أن تدون البيانات التالية عن السائح .

١. الاسم ، الجنس ، الجنسية والمهنة ، محل الميلاد .
٢. بلد الإقامة الدائم .
٣. رقم جواز السفر وتاريخ ومكان إصداره
٤. طريقة القدوم (جوا – برا – بحرا)
٥. تاريخ الوصول
٦. جهة القدوم
٧. مكان الإقامة (فندق – قرية سياحية – مسكن خاص – لدى الأصدقاء – أخرى تذكر)
٨. تاريخ المغادرة

ثانيا : تسجيل البيانات الإحصائية السياحية عند النزول : -

يتحقق ذلك فى الدفاتر الممسوكة (يدويا أو بالكمبيوتر) فى الفنادق على أن تدون البيانات التالية عن السائح : -

١. اسم الفندق ومكان وجوده
٢. اسم النزيل والحالة الاجتماعية

٣.الجنسية ومكان وتاريخ الميلاد

٤.مكان الإقامة الدائم

٥.رقم جواز السفر وتاريخ ومكان صدوره .

٦.تاريخ الوصول

٧.تاريخ المغادرة .

ثالثا : تسجيل البيانات الإحصائية السياحية عند المغادرة :-

ويتحقق ذلك فى الدفاتر الممسوكة (يدويا أو على الكومبيوتر) على الحدود والمنافذ الشرعية للدولة على أن تدون بيانات السائحين عند العودة إلى بلادهم وانه فى حالة ما إذا كان البحث يستهدف جمع معلومات وبيانات إحصائية عن السائحين بخلاف المعلومات والبيانات السابقة فان الأمر يستلزم عمل استمارة استبيان ، وكما سبق القول فان إعداد استمارة الاستبيان يتوقف على الهدف من البحث ، فقد يكون المطلوب هو تصميم نموذج لاستمارة استبيان يستهدف دراسة النمط الاستهلاكي للسائحين العرب .



الفصل الثالث

الحصر الشامل والعينة في جمع البيانات

عند جمع البيانات قد يقوم الباحث بجمع البيانات من جميع مفردات المجتمع الأصلي محل الدراسة وهذا ما يعرف بالحصر الشامل ، لكن في الواقع التطبيقي غالبا يتم اختيار جزء من مفردات المجتمع الإحصائي محل الدراسة للقيام بجمع البيانات من هذا الجزء فقط ، هذا الجزء هو ما يطلق عليه اسم عينه Sample .

ويفضل استخدام أسلوب العينة في جمع البيانات الإحصائية للأسباب التالية :-

١. عدم ملائمة المجتمع الإحصائي محل الدراسة لإجراء الحصر الشامل لجميع مفرداته، فمثلا إذا أريد إجراء دراسة لمعرفة متوسط وزن السمكة في البحر الأحمر فلا يمكن اصطياد جميع الأسماك الموجودة في البحر لإجراء تقدير متوسط وزن السمكة ، لكن يمكن إجراء هذا التقدير باستخدام عينة مأخوذة من سمك هذا البحر .

٢. عدم كفاية الموارد المادية والفنية المتاحة لإجراء البحث .

٣. ضيق الفترة الزمنية المعطاة لإجراء البحث .

٤. في بعض الأحيان تعطى العينة نتائج اقرب إلى الصحة

من نتائج الحصر الشامل ، فبالرغم من أن المجتمع الإحصائي قد يناسبه أسلوب الحصر الشامل في جمع البيانات إلا أن الباحث قد يرى أن مفردات هذا المجتمع سوف يدلون ببيانات خاطئة معتقدين أن إجاباتهم هذه ستعود عليهم بالنفع ، وهنا يضطر الباحث إلى اختيار مفردات بعينها كعينة من هذا المجتمع الإحصائي .

٥. إذا اتصف المجتمع الإحصائي محل الدراسة بالتجانس فان جمع البيانات من هذا المجتمع بأسلوب الحصر الشامل ما هو إلا مضيعة للوقت والجهد إذ يكفى جمع البيانات من عينة مسحوبة منه والتوصل منها إلى نتائج لن تختلف اختلافا يذكر عن نتائج الحصر الشامل ، فمثلا فحص جودة قصاصة صغيرة من قطعة قماش كبيرة تكفى لشراء هذه القطعة إذا ثبت الفحص جودة القصاصة ، وكذلك الحفنة المأخوذة من جوال أرز تكفى للحكم على نوعية هذا الأرز ، وأيضا أنبوبة صغيرة من ماء البحر تكفى لدراسة تركيب ماء البحر كله وهكذا.

ومن الأهمية بمكان الإشارة إلى ما يجب أن يراعيه الباحث قبل الشروع فى اخذ العينة من المجتمع الإحصائي محل الدراسة : -

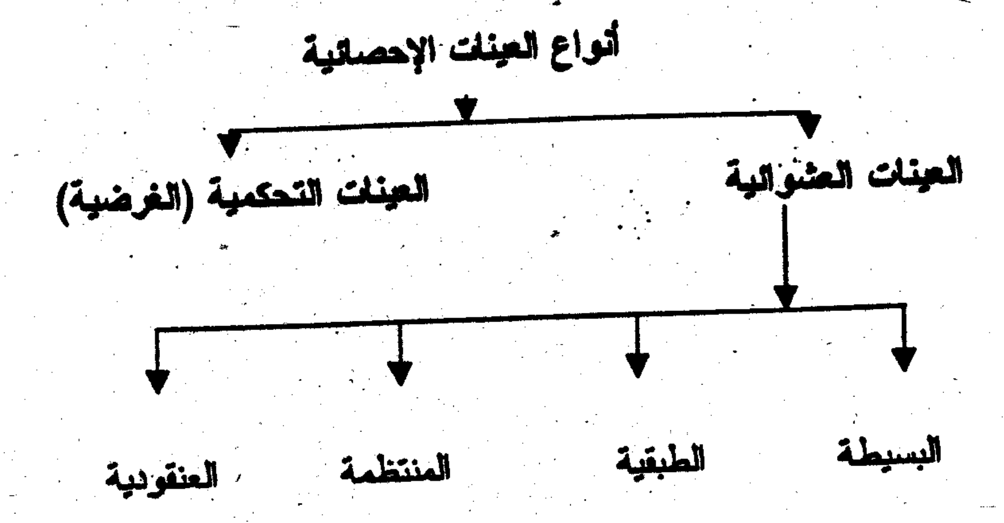
١. أن يقوم بعملية التحديد للمجتمع الإحصائي محل الدراسة بمعنى تحديد مجموعة المفردات التى يتكون منها هذا المجتمع والتى يراد دراستها لمعرفة خصائص هذا المجتمع بناءا على دراسة العينة المسحوبة منه ، ويطلق على مجموعة المفردات هذه



اسم إطار المعاينة Sampling Frame ، ويعتبر إعداد الإطار أمر ضرورى لاختيار العينة ، ويجب أن يكون الإطار جيدا بحيث يشمل جميع مفردات المجتمع الإحصائي محل الدراسة دون تكرار ، ولا تكون العينة ممثلة للمجتمع إذا كان هناك اختلاف بين الإطار والمجتمع المستهدف دراسته ، فمثلا لا يجوز استخدام مجموعة السائحين القادمين إلى مصر فى سنة ما كإطار يتم اختيار عينة منه لدراسة النمط الاستهلاكي للسائحين العرب .

٢. تحديد المتغيرات الإحصائية المقصودة والتى يتم جمع بياناتها من المفردات المأخوذة فى العينة .

٣. الاختيار الصحيح لنوع العينة المناسب لإجراء الدراسة حيث تتعدد أنواع العينات الإحصائية .



أولا : العينة العشوائية البسيطة Simple Random Sample

هى عينة يتم اختيارها من بين كل العينات الإحصائية التى يتكون منها المجتمع الإحصائي محل الدراسة ، بشرط أن يعطى لكل



من هذه العينات فرص متساوية فى أن يتم اختيار آيا منها ، حيث أن جميع مفردات المجتمع متجانسة وبالتالى لا يوجد ما يدعو لتفضيل عينة عن الأخرى .

والسؤال الآن كيف يمكن معرفة العدد الكلى للعينات الإحصائية المكونة للمجتمع الإحصائي وأيضا معرفة عناصر كل منها؟ والإجابة تستلزم أن يتم التناول للمثال التالى مع الاستعانة بالبعد الرياضى .

المثال

نفرض أن مجتمع إحصائي يتكون من ٦ أفراد هم احمد ، بهجت ، جابر ، دولت ، هناء ، وفاء وانه يراد اختيار فردين (مجوود فردين) كعينة من هذا المجتمع . فما هو العدد الكلى للعينات المكونة لهذا المجتمع ؟ وما هى عناصر كل منها ؟

الحل

١) نفرض أن الرموز أ ، ب ، جـ ، د ، هـ ، و هى رموز بديلة عن الأسماء المذكورة أي أن عناصر المجموعة الشاملة س هى { أ ، ب ، جـ ، د ، هـ ، و } .

٢) وباستخدام الأسس يتم معرفة العدد الكلى للأزواج المرتبة التى يمكن تكوينها من عناصر المجموعة الشاملة ، ولما كانت عناصر المجموعة الشاملة هى ٦ والزوج المرتب يتكون من عنصرين ، إذاً

$٦^{٢}$ = ٣٦ زوج مرتب .



٣) ويتم معرفة عناصر كل زوج مرتب كما يلى : -

{أ ، ب ، جـ ، د ، ، هـ ، و}

= { (أ ، أ) ، (أ ، ب) ،(أ ، جـ)، (أ ، د)،(أ ، هـ)، (أ ، و)

،(ب ، أ)، (ب ، ب)،(ب ، جـ)، (ب ، د)،(ب ، هـ)،(ب ، و)

،(جـ، أ)،(جـ ، ب)،(جـ ، جـ)،(جـ ، د)،(جـ ، هـ)،(جـ ، و)

،(د ، أ)، (د ، ب)،(د ، جـ)،(د ، د)،(د ، هـ)،(د ، و)

،(هـ ، أ)،(هـ ، ب)،(هـ ، جـ)،(هـ ، د)،(هـ ، هـ)،(هـ ، و)

،(و ، أ)،(و ، ب)،(و ، جـ)،(و ، د)،(و ، هـ)،(و ، و) }

٤) ولما كان المفكوك السابق يشتمل على أزواج مرتبة عناصرها مكرره مثل (أ ، أ) ، (ب ، ب) ، ... فلا يمكن أن يطلق على تلك الأزواج المرتبة اسم عينة Sample ، لذلك يتم استبعاد تلك الأزواج المرتبة من المفكوك السابق ليصبح المفكوك فى الوضع الجديد كما يلى : -

{أ ، ب ، جـ ، د ، ، هـ ، و}

= { (،)،(أ ، ب)،(أ ، جـ)،(أ ، د)،(أ ، هـ)،(أ ، و)

،(ب ، أ)،(،)،(ب ، جـ)،(ب ، د)،(ب ، هـ)،(ب ، و)

،(جـ ، أ)،(جـ ، ب)،(،)،(جـ ، د)،(جـ ، هـ)،(جـ ، و)

،(د ، أ)،(د ، ب)،(د ، جـ)،(،)،(د ، هـ)،(د ، و)

،(هـ ، أ)،(هـ ، ب)،(هـ ، جـ)،(هـ ، د)،(،)،(هـ ، و)

،(و ، أ)، (و ، ب)،(و ، جـ)،(و ، د)،(و ، هـ)،(،) }

وبالتالى يصبح العدد الكلى للأزواج المرتبة فى الوضع الجديد هى ٣٠ زوج مرتب ، وباستخدام قانون التبادل يمكن معرفـــة العــدد الكلـــى للأزواج المرتبة فى شكل تراتيب وذلك قبل إجراء عملية الفـــك كمـــا يلــــى: -

$$^{\text{ن}}\text{ل}_{\text{ر}} = \frac{\underline{|\text{ن}}}{\underline{|\text{ن} - \text{س}}}$$

حيث :

ن : عدد الأشياء المراد ترتيبها ر من الترتيبات(عدد ترتيبات)

ر : عدد عناصر كل ترتيبه

$$^{٦}\text{ل}_{٢} = \frac{\underline{|٦}}{\underline{|٦ - ٢}} = \frac{\underline{|٦}}{\underline{|٤}}$$

$$= \frac{١×٢×٣×٤× ٥×٦}{١×٢×٣×٤}$$

= ٣٠ ترتيبه

٥) ولما كان المفكوك الجديد السابق يشتمل على أزواج مرتبة متشابهة مثل (أ ، ب) ، (ب ، أ) ، (ا ، جـ) ، (جـــ ، أ) ، .... فانه يكتفى بواحدة منعا للتكرار ، وعليه يصبح المفكوك الجديد فـــى الوضع الجديد الآخر كما يلى : -

{أ ، ب ، جـ ، د ، هـ ، و} =

= { (،)، (أ ، ب)، (أ ، جـ)، (أ ، د)، (أ ، هـ)، (أ ، و)

، (،)، (،)، (ب ، جـ)، (ب ، د)، (ب ، هـ)، (ب ، و)

، (،)، (،)، (،)، (جـ ، د)، (جـ ، هـ)، (جـ ، و)

، (،)، (،)، (،)، (،)، (د ، هـ)، (د ، و)

، (،)، (،)، (،)، (،)، (،)، (هـ ، و)

، (،)، (،)، (،)، (،)، (،)، (،) }

وباستخدام قانون التوافين يمكن معرفة العدد الكلى للأزواج المرتبة فى شكل توافيق (عينات) وذلك قبل إجراء عملية الفك كما يلى : -

$$ {}^{ن}ق_{ر} = \frac{\underline{|ن}}{\underline{|ر}\ \underline{|ن - س}} $$

$$ = \frac{\underline{|٦}}{\underline{|٢}\ \underline{|٤}} $$

$$ = \frac{١ \times ٢ \times ٣ \times ٤ \times ٥ \times ٦}{١ \times ٢ \times ٣ \times ٤ \times ١ \times ٢} $$

= ١٥ توفيقه (عينة)

وهذا الرقم هو بالفعل عدد الأزواج المرتبة فى شكل عينات بالمفكوك السابق .

ويتضح مما سبق أننا أمام ١٥ عينة أمكن تكوينها من الستة أفـراد بحيث كل عينة منهم تتكون من فردين ، والعينات هى : –

(احمد ، بهجت) ، (احمد ، جابـر) ، (احمد ، دولـت)
، (احمد ، هنـاء) ، (احمد ، وفـاء) ، (بهجت ، جابـر)
، (بهجت ، دولت) ، (بهجت ، وهناء) ، (بهجت ، وفاء)
، (جابـر ، دولت) ، (جابـر ، هناء) ، (جابر ، وفـاء)
، (دولت ، هنـاء) ، (دولت ، وفـاء) ، (هنـاء ، وفـاء).

و الآن كيف يمكن اختيار عينة واحدة مـن تلـك العينـات اختيارا عشوائيا ؟

الحـل

أ – الاختيار باستخدام البطاقات (المعاينة بالبطاقات)
Tichet Sampling

فى هذه الطريقة يتم كتابة اسم كل عينة من العينات الخمسـة عشر فى المثال السابق على ١٥ قطعة من الورق بمعنى أن لكل عينة ورقة مستقلة بشرط أن هذا الورق متساوى ومتماثل من حيث الشـكل والحجم ، ثم يتم وضع قطع الورق هذه بعد طيها فى صندوق ، وبعـد الخلط الجيد لتلك الأوراق فى الصندوق يتم سحب ورقة واحدة وتدوين ما هو مكتوب عليها ، ولنفرض أنها (هناء ووفاء) فتكون هذه هـى عينة عشوائية بسيطة .



ب- الاختيار باستخدام جداول الأعداد العشوائية :-

يتم اللجوء إلى جداول الأعداد العشوائية Table of Random Numbers إذا كان عدد مفردات المجتمع الإحصائي كبير مما يصعب معه استعمال الطريقة السابقة في الاختيار الصحيح لمفردات العينة من المجتمع .

وجداول الأعداد العشوائية عبارة عن مجموعة الأرقام من صفر إلى ٩ مرتبة عشوائيا فى شكل صفوف و أعمدة وعلى هيئة مجموعات كبيرة . فإذا كنا بصدد اختيار عينة حجمها ٣٠ مفردة من بين ٩٩ مفردة فانه يتم الآتى :

- ترقيم المفردات ٩٩ بالأرقام ١ ، ٢ ، ٣ ، ..... ، ٩٩ حيث يصبح لكل مفردة رقم .
- استخدام جدول الأعداد العشوائية ذو الرقمين .
- اختيار رقم ما عشوائيا كنقطة بداية من أرقام أحد أعمدة جدول الأعداد العشوائية .
- الاستمرار فى اختيار الأرقام العشوائية بعد نقطة البداية بشرط السير فى أي اتجاه لكن بشكل منظم ، واستبعاد الأعداد المكررة ، وأيضا استبعاد الأعداد التى تزيد عن ٩٩، وذلك حتى يتم اختيار عدد المفردات ٣٠ .

ثانيا : العينة العشوائية الطبقية

Stratified Random Sample

يتم استخدام هذه العينة إذا كانت مفردات المجتمع الإحصائي محل الدراسة بينها اختلافات واضحة ، فهنا يتم تقسيم هذا المجتمع إلى



طبقات Strata بحيث تكون الفروق صغيرة داخل كل طبقة ، بينما تكون الفروق كبيرة نسبيا بين الطبقات . ثم بعد ذلك يتم اختيار مفردات العينة العشوائية الطبقية من هذا المجتمع على أي من الأساسين :

الأول : أن نسبة المفردات فى طبقات العينة تكون كنسبة المفردات فى طبقات المجتمع .

الثانى : إذا كانت طبقات المجتمع بعضها متجانس و الآخر غير متجانس { يوجد تشتت كبير بين قيم مفرداتها } ، فانه يتم اختيار عدد مفردات قليل من الطبقات المتجانسة وعدد مفردات كبير من الطبقات غير المتجانسة .

مع مراعاة أن يكون اختيار مفردات كل طبقة بالعينة بالطريقة العشوائية البسيطة السابق ذكرها .

مثال

إذا كان عدد الطلبة البنين بالسياحة والفنادق هو ١٥٠٠ طالب، وعدد الطلبة البنات هو ٥٠٠ طالبة . والمطلوب سحب عينة عشوائية طبقية بحجم٢٠٠مفردة من هذا المجتمع لدراسة أطوال الطلاب وذلك :

- بفرض أن طبقتى البنين والبنات فى المجتمع تجانسهما متقارب .

- بفرض أن طبقة البنات فى المجتمع تجانسها اكبر من تجانس طبقة البنين (يعنى أن قيم مفردات طبقة البنين تشتتها اكبر).

- بفرض تقارب التجانس في

هنا تكون نسبة الطبقتين فى

، .. نسبة الطبقتين فى ا

= ١٥٠٠ :

= ٣ : ١

.. نسبة الطبقتين فى العينة

= ٣ : ١

.. عدد مفردات طبقة البنين فى العينة

$= \frac{٣}{٤} \times ٢٠٠$

= ١٥٠ طالب

، عدد مفردات طبقة البنات فى العين

$= \frac{١}{٤} \times ٢٠٠$

= ٥٠ طالبة

.. العينة العشوائية الطبقية بحجم ٢٠٠ مفردة من هـذا المجتمـع تتكون من ١٥٠ طالب ، ٥٠ طالبـة يتـم سـحبها بالطريقـة

العشوائية البسيطة .

- بفرض تجانس طبقة البنات وعدم تجانس طبقة البنين :

هنا يتم سحب عدد مفردات قليل من طبقة البنات وليكن ٣٥ طالبة ، وسحب عدد مفردات كبير من طبقة البنين وليكن ١٦٥ طالب، وذلك لضمان تمثيل العينة للمجتمع .

ثالثا : العينة العشوائية المنتظمة

Systematic Random Sample

تستخدم هذه العينة إذا كان المجتمع محل الدراسة مفرداته مرتبة ، مثل كشوف أسماء طلبة قسم الفنادق المرتبة حسب مجموع الدرجات ، أو كشوف أرقام المنازل المرتبة رقميا بأحد الأحياء ،أو كشوف العاملين بأحد الشركات السياحية المرتبة حسب قيم الأجور ، ويتم اختيار مفردات العينة العشوائية المنتظمة من مجتمعها محل الدراسة كما يلي :

تحديد رقم معدل الاختيار وهو = $\frac{\text{حجم المجتمع}}{\text{حجم العينة}}$

- استخدام رقم معدل الاختيار في تكوين عدد مفردات المجموعة الأولى .
- من مفردات المجموعة الأولى هذه يتم اختيار مفردة عشوائيا لتكون هي المفردة الأولى في العينة العشوائية المنتظمة .
- يتم اختيار المفردة الثانية بإضافة رقم معدل الاختيار إلى رقم



المفردة الاولى ، وهكذا يستمر نفس الأسلوب حتى يتم اختيار عـدد مفردات العينة العشوائية المنتظمة .

مثال

مجتمع إحصائي ما يحتوى إطاره على ٥٠٠٠ مفردة مرتبـة والمطلوب سحب عينة عشوائية منتظمة بحجم ٥٠٠ مفردة مـن هـذا المجتمع .

الحـل

$$\text{رقم معدل الاختيار} = \frac{\text{حجم المجتمع}}{\text{حجم العينة}}$$

$$= \frac{٥٠٠٠}{٥٠٠٠}$$

$$= ١٠$$

.. مفردات المجموعة الأولى هى :

١ ، ٢ ، ٣ ، ٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ ، ١٠

.. المفردة الأولى فى العينة العشوائية المنتظمة هى ٦ (تم سحبها من مفردات المجموعة الأولى بأسلوب العينة العشوائية البسيطة) .

.. المفردة الثانية فى العينة العشوائية المنتظمة هى :

٦ + ١٠ = ١٦



∴ المفردة الثالثة فى العينة العشوائية المنتظمة هى :

١٦ + ١٠ = ٢٦

∴ مفردات العينة العشوائية المنتظمة بحجم ٥٠٠ مفردة والمسحوبة من المجتمع ذو ٥٠٠٠ مفردة هى المفردات التى تحمل الارقام ٦ ، ١٦ ، ٢٦ ، ٣٦ ، ٤٦ ، .... ، ٤٩٩٦.*

تمتاز هذه العينة بالسهولة ، كما أنها تمثل المجتمع كله بطريقة متساوية ، لكن يعاب عليها فى عدم صلاحيتها إذا كانت المفردات المرتبة فى المجتمع بينها علاقة دورية وكان معدل الاختيار لمفردات العينة مساويا لطول الدورة أو لإحدى مضاعفاتها .

رابعا : العينة العشوائية متعددة المراحل

Malti-Stage Random Sample

تستخدم هذه العينة إذا كان المجتمع الإحصائي محل الدراسة بطبيعته مقسم إلى قطاعات ، ولذلك يتم الاختيار العشوائى لمفردات هذه العينة على اكثر من مرحلة ، فمثلا يتم الاختيار العشوائى لبعض المراكز فى كل محافظات ج.م.ع ، ثم بعض القرى فى كل مركز مختار ، ثم بعض الأسر من مجموع أسر كل قرية مختارة ، فنحصل على المفردات (الأسر) المطلوبة للعينة .

خامسا : العينة المقصودة (الغرضية) : Purposive Sample

تستخدم هذه العينة إذا كان المجتمع الإحصائي محل الدراسة لا

---

* هذا الرقم جاء من قانون الحد الأخير فى المتوالية العديدية وهو :

ح = أ + د (ن – ١) .



يسمح إلا باختيار عدد قليل من مفرداته ، وبالتـــالى فشـــروط العينـــة العشوائية غير متوفرة وهى إعطاء جميع مفردات المجتمـــع فـــرص متساوية فى الاختيار ، لذلك يصبح لا مفر من استخدام هذه المفـــردات القلائل بطريقة يعتقد الباحث فيها توافر صفر تمثيل هـــذه المفـــردات للمجتمع ، ومن الواضح أن هذا النوع من العينات لا يستبعد اثر التحيز الشخصى .

وجدير بالذكر أن اختيار نوع العينة يتوقف على نوع الظاهرة وعلى طبيعة المفردات محل الدراسة .



الفصل الرابع

الأخطاء الشائعة
عند جمع البيانات الإحصائية

تكون البيانات الإحصائية خاطئة إذا ما تعرضت لخطأ أو لآخر من الأخطاء التالية : -

١. الخطأ الذى ينشأ عند نقل أو نشر البيانات .

٢. الخطأ الذى ينشأ عن عدم توافق التعاريف المستخدمة فى البحث محل الدراسة مع التعاريف المستخدمة فى البيانات المسجلة فى الجهات الرسمية ، فقد يكون البحث محل الدراسة عن أهم صادرات ج.م.ع فى حين أن البيانات المسجلة عن إجمالي صادرات ج . م . ع .

٣. الخطأ الذى ينشأ من عدم التحديد الدقيق للبيانات المطلوبة بسبب الفشل فى تحديد فروض المشكلة البحثية ، فقد تكون المشكلة البحثية هى تراجع أعداد السائحين العرب إلى مصر ويضع الباحث فرضا مؤداه أن عدم جودة لحوم الخنازير المقدمة هى سبب هذه المشكلة ، مثل هذا الفرض غير الدقيق يجعل البيانات التى يتم جمعها ببيانات خاطئة .

٤. الخطأ الذى ينشأ من عدم تحديد المجتمع الإحصائي محل الدراسة مما يترتب عليه استخدام إطار معيب لهذا المجتمع فيتسبب فى اختيار مفردات يعتقد أنها تمثل المجتمع وهى لا تمثله ، فاستخدام أصحاب التليفونات المحمولة كإطار لدراسة مشكلة تخص المجتمع ككل يعتبر إطار معيب حيث أغفل باقى المواطنين الذين تخصهم نفس المشكلة ولا يمتلكون تليفونات محمولة ، ومن ثم فالبيانات التي يتم جمعها فى هذه الحالة بيانات خاطئة .

٥. الخطأ الذى ينشأ عن الإجابات الخاطئة التى يدلى بها المبحوثين فى استمارة الاستبيان سواء عن سهو أو خطأ أو تعمد .

٦. الخطأ الفنى الذى يقع فيه الباحث نفسه عند تبويب وتصنيف البيانات و إعداد الجداول واختيار المقاييس المناسبة .

٧. الخطأ العشوائى : -

يسمى الخطأ العشوائى Rondom Error بخطأ الصدفة Chance Error وهو الخطأ الذى ينشأ عند تقدير معلمة المجتمع باستخدام عينة عشوائية واحدة من بين عدد العينات الممكنة فى المجتمع .

وبديهى القول انه لا يوجد خطأ عشوائى إذا تم تقدير معالم المجتمع باستخدام أسلوب الحصر الشامل ، ولا يعنى خلو أسلوب الحصر الشامل من الخطأ العشوائى بأنه الأسلوب الأفضل فى جمع البيانات عن أسلوب العينة ، بل على العكس حيث يكبر مجموع الأخطاء غير العشوائية عند جمع البيانات من جميع المفردات أي الحصر الشامل ، بينما تقل هذه الأخطاء فى حالة العينة بسبب تمكن



جامع البيانات من الحصول على بيانات دقيقة من المفـــردات القليلـــة (العينة) بدون جهد كبير بخلاف الحال عند جمع البيانـــــات بأســلوب الحصر الشامل . وحتى الأخطاء العشوائية المأخوذة علـــى أسـلوب العينة فانه يمكن استخدام العينـــة المناســبة والأســاليب الرياضيـــة المستخدمة في إحصاء العينات لقياس هذا الخطأ وتقديره و تدنيته إلــى حد يقبله الباحث مقدما *

٨.خطأ التحيز : Bias

هو خطأ ينشأ من تصرفات القائمين بالبحث أنفسهم عند جمـــع البيانات من عينة ، ومن هذه التصرفات : -

أ - إعطاء مفردات المجتمع فرص غير متساوية عند اختيارها في العينة .

ب- جمع البيانات من مفردات لا تشملها العينـــة ، أو إغفـــال بعض مفردات العينة عند جمع البيانات .

جـ - قد تكون البيانات التى تم جمعها من العينة سليمة ، لكن قد يحدث خطأ عند حساب معالم المجتمع من نتائج العينة، مثال ذلك إذا كنا بصدد تقدير متوسط دخل الأســـرة فـــى مدينة ما ، وتم اخذ عينتين مختلفتي الحجم الأولى بحجـــم ٣٠٠ أسرة والثانية بحجم ١٥٠ أسرة ، وبعد جمع بيانات الدخل من تلك الأسر تبين أن متوسط دخل الأســـرة فـــى العينة الأولى ٣٠ ألف جنية وان متوسط دخل الأسرة فـــى

---

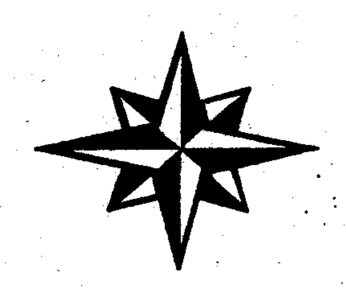
* سوف يتم التعرف على ذلك عند دراسة التقدير الإحصائي لمعالم المجتمع من عينة فى الجزء الثانى من هذا الكتاب .

العينة الثانية ٢٠ ألف جنية ، فيكون من الخطأ حساب متوسط دخل الأسرة في المدينة (المجتمع) على أساس متوسط متوسطى العينتين أي $\frac{٣٠ + ٢٠}{٢}$=٢٥ ألف جنية، و إنما يتم الحساب على أساس الترجيح بالأوزان أي

$$\frac{١٥٠×٢٠ + ٣٠٠×٣٠}{١٥٠ + ٣٠٠}$$ = ٢٦,٧ ألف جنية .

ملحوظة :-

يتضح مما سبق أن بيانات العينة تشتمل على نوعين من الخطأ هما خطأ الصدفة وخطأ التحيز ، ويسميان معا بخطأ المعاينة . وخطأ المعاينة هو الفرق بين (إحصاء العينة) و (نظيرتها من معالم المجتمع المطلوب تقديرها) . وانه إذا بذل الجهد فى تخليص العينة من خطأ التحيز فإنها لا تسلم من خطأ الصدفة والذى يعد أمرا محتوما ألا انه فى الإمكان حساب حدود احتمالاته وسوف نرى ذلك عند دراسة الاحتمالات وإحصاء العينات .



الباب الثالث

الدراسة الاحصائية لمتغير احصائي واحد

الدراسة الإحصائية لمتغير إحصائي واحد

يشتمل هذا الباب على الفصول التالية : -

الفصل الأول : عرض البيانات الإحصائية

- في حالة المجتمع الإحصائي الصغير
- في حالة المجتمع الإحصائي الكبير

الفصل الثاني : النزعة المركزية للبيانات الإحصائية

- في حالة المجتمع الإحصائي الصغير
- في حالة المجتمع الإحصائي الكبير

الفصل الثالث : تشتت البيانات الإحصائية

- في حالة المجتمع الإحصائي الصغير
- في حالة المجتمع الإحصائي الكبير



الفصل الأول

عرض البيانات الإحصائية

مقدمه :-

بعد الانتهاء من عملية جمع البيانات الإحصائية لابد من إجراء مراجعة على هذه البيانات ، وذلك لكشف أي أخطاء بها وللتأكد من صحتها .

وتتم عملية المراجعة من خلال التأكد من أن الأمئلة في استمارة الاستبيان قد تم الإجابة عليها كاملة وتستبعد التى لا تنطبق عليها هذا الوصف ، كما يتم التأكد من أن الإجابات معقولة وغير متناقضة مع بعضها ، فإذا كان هناك تناقض أو إجابات غير محتملة الصحة فيجب تصحيحها أو ردها إلى مصادرها الأصلية لتصحيحها أو أن يتم استبعادها، كما يجب التحقيق من تجانس الوحدات المستخدمة فى الإجابة فمثلا إجابة سؤال عن الأجر فقد يجيب المستقصى منه فى استمارة الاستبيان عن الأجر فى السنة أو فى الشهر أو فى اليوم لذلك يجب توحيد الأساس لجميع الأجور .

بعد ذلك يتم عرض البيانات الإحصائية بأي من طرق العرض الثلاثة التالية :



١- عرض البيانات فى صيغة كتابية : -

تعد طريقة الصيغة الكتابية فى عرض البيانات الإحصائية من ابسط الطرق حيث أن البيانات المطلوب عرضها يتم دمجها فى سياق الكلام . وهذه الطريقة ضرورية إذا كانت البيانات المطلوب عرضها قليلة ولا يكون من المناسب عرضها فى صورة جدول ، وترجع أهمية هذه الطريقة أن عرض الأرقام فى سياق الكلام يلفت نظر القارئ إليها .

إلا أن يعاب على هذه الطريقة بأنها مطولة وعقيمة وتجعل الملل يتسرب إلى القارئ لأنه يضطر إلى قراءة النص المكتوب كله للتعرف على البيانات المراد عرضها مما يتطلب منه وقت طويل .

مثال

إذا ما تجمعت بيانات عن الفنادق فى ج.م.ع فانه يمكن عرض تلك البيانات فى صورة تقرير كتابى كما يلى : -

انه فى عام ١٩٩٠ كان إجمالي عدد غرف الفنادق فى ج.م.ع يبلغ ٥١٨٠٩ غرفة خص الفنادق الثابته ٣٦١١٢ غرفة والفنادق العائمة ٧٨٢٥ غرفة والقرى السياحية ٧٨٧٢ غرفة ، وفى عام ١٩٩١ كان إجمالي عدد غرف الفنادق فى ج.م.ع يبلغ ٥٣٧٢٧ غرفة خص الفنادق الثابتة ٣٧١٢٠ غرفة والفنادق العائمة ٨٨٣٠ غرفة والقرى السياحية ٧٧٧٧ غرفة ، وفى عام ١٩٩٢ كان إجمالي عدد غرف الفنادق فى ج.م.ع يبلغ ٥٥٦١٠ غرفة خص الفنادق الثابتة ٣٧٤٩٤ غرفة والفنادق العائمة ٩٢٩٧غرفة والقرى السياحية ٨٨١٩، وفى عام ١٩٩٣ كان إجمالي عدد غرف الفنادق فى ج.م.ع يبلغ ٥٨٧٥٥ غرفة خص



الفنادق الثابتة ٣٩٤٤١ غرفة والفنادق العائمة ٩٧٦٣غرفة والقرى السياحية ٩٥٥١ غرفة ، وفى عام ١٩٩٤ كان إجمالي عدد غرف الفنادق فى ج. م . ع يبلغ ٦١٠٦٨ غرفة خص الفنادق الثابتة ٤٠٣٨٠غرفة والفنادق العائمة ١٠٣٣٩غرفة والقرى السياحية ١٠٣٤٩ غرفة ، وفى عام ١٩٩٥ كان إجمالي عدد غرف الفنادق فى ج. م . ع يبلغ ٦٤٩٥٨ غرفة خص الفنادق الثابتة ٤٣٧٢٨غرفة والفنادق العائمة ١٠٥٣٢غرفة والقرى السياحية ١٠٦٩٨ غرفة ، وفى عام ١٩٩٦ كان إجمالي عدد غرف الفنادق فى ج. م . ع يبلغ ٧٠٤٧١ غرفة خص الفنادق الثابتة ٤٧٥٧٣ غرفة والفنادق العائمة ١١١٨٤غرفة والقرى السياحية ١١٧١٤ غرفة ، وفى عام ١٩٩٧ كان إجمالي عدد غرف الفنادق فى ج. م . ع يبلغ ٧٥٦٧٩ غرفة خص الفنادق الثابتة ٥٠٢٣٩ غرفة والفنادق العائمة ١١٣٢٢غرفة والقرى السياحية١٤١١٨ غرفة ، وفى عام ١٩٩٨ كان إجمالي عدد غرف الفنادق فى ج. م . ع يبلغ ٨٢٩٢٥غرفة خص الفنادق الثابتة ٥٢٤٦٨ غرفة والفنادق العائمة ١١٦٧٣غرفة والقرى السياحية ١٨٧٨٤ غرفة .

٢- عرض البيانات فى صورة جدولية : -

لما كان عرض البيانات الإحصائية ضمن سياق الكلام أمر لا يتيح للقارئ استيعاب تلك البيانات ولا يمكن من الالمام بمدلولاتها ، لذلك يلجا الباحثون إلى عرض بياناتهم فى جداول إحصائية .

والجدول الإحصائي Statistical table هو عبارة عن ترتيب منظم للبيانات فى صورة صفوف و أعمدة بقصد إيجاز البيانات مع ضرورة الإيضاح والسهولة فى قراءتها وفهم مضمونها .



وهناك عدة نقاط يجب مراعاتها عند عمل الجدول الإحصائي اللازم وهى : -

١. رقم الجدول : Number

يجب أن يعطى للجدول رقم معين حتى يسهل الرجوع إليه عند الحاجة .

٢.عنوان الجدول : Title

لابد أن يحمل الجدول الإحصائي عنوانا مختصرا وفى نفس الوقت مفسرا لمحتوياته فيتضمن : -

- نوع البيانات What مثل عدد السائحين
- مكان جمع البيانات Where مثل عدد السائحين بجمهورية مصر العربية .
- أساس تقسيم البيانات How فيقال عدد السائحين فى ج.م.ع حسب الجنسية بالألف
- فترة جمع البيانات When فيقال عدد السائحين فى ج.م.ع حسب الجنسية بالألف خلال الفترة (١٩٩٠ – ١٩٩٥) .

٣.الوحدات المستعملة :

يجب تحديد الوحدات المستعملة فى البيانات فتكون بالطن أو الكيلوجرام أو بالجنيهات أو القرش أو بالدولار ....الخ .

٤.حجم الجدول وشكله العام وتقسيماته : -

من المرغوب فيه دائما أن يكون الجدول متناسقا أي تتناسب أعمدته مع صفوفه ، فلا يكون طويلا وضيقا أو قصيرا ومتسعا ، كما يجب أن تكون تقسيماته منسقة وسهلة الفهم وواضحة ، وعادة يفضل الاحصائيون تقليل التسطير في الجدول ، فالتسطير الأفقي غير مرغوب والتسطير الرأسي يفضل إلا يكون كثيرا ، كما يفضل أن تكون طريقة كتابة الجدول غير مجهده للنظر فنترك مثلا فراغات بين مجموعات الأرقام ، والا تكون متلاصقة إلى جانب بعضها .

٥. المذكرة التفسيرية السفلية: **Foot notes**

قد تتطلب بعض بيانات الجدول إيضاحات تفسيرية ، وهذه تكتب عادة اسفل الجدول مباشرة ، وتلك المذكرات عادة تفسر بيان خاص أو رقم معين فى صلب الجدول . وجرت العادة على أن يوضع فوق هذا الرقم أو البيان علامة معينة (* مثلا) ، وهذه تظهر أن هناك تفسيرا لها فى نهاية الجدول .

٦. المصدر : **Source**

يتم كتابة مصدر البيانات اسفل المذكرة التفسيرية ؛ وذكر المصدر أمر ضرورى لأنه يزيد من ثقة القارئ بهذه البيانات ، وأيضا لإمكان الرجوع إليه عند الرغبة فى بيانات اكثر تفصيلا ، أو عند الرغبة فى التأكد من رقم معين أو إيجاد تفسير له .

٧. ترتيب البيانات فى الجدول : **Arrangement of items**

ترجع أهمية ترتيب البيانات داخل الجدول إلى أن هذا الترتيب يساعد على سهولة فهم بيانات الجدول وأيضا إلى سهولة تحليل نتائجه



بالإضافة إلى إمكانية إجراء المقارنة بين بيانات الجدول بسهولة ، ومن طرق ترتيب البيانات داخل الجدول : -

أ – الترتيب الابجدى : Alphabetical

وهذا الترتيب يتم استعماله عادة فى الجداول العامه وذلك لسهولة تحديد مكان أي رقم بالجدول .

ب– الترتيب حسب القيمة : Magnitude

وهو ترتيب للبيانات داخل الجدول على أساس كتابة اكبر الأرقام أولا سواء فى السطر أو العمود ثم يلى الرقم الأكبر الرقم الأصغر فالأصغر .

جـ – الترتيب حسب الموقع الجغرافى : Geographical

وفيه يتم كتابة البيانات داخل الجدول حسب موقعها الجغرافى كأن يكتب اسم المحافظة ثم يليها اسم الحى ثم اسم الشارع .

ء – الترتيب حسب التقسيم الدارج :

وفيه يتم ترتيب البيانات داخل الجدول وفقا لما هو متفق عليه بين الباحثين .

والجدول التالى يمكن اعتباره نموذجا لجدول إحصائي روعى فى كتابته النقاط السابقة : -

جدول رقم (   ) تطور الطاقة الفندقية المصرية خلال الفترة (٩٣ - ١٩٩٩)

(العدد بالوحدة)

السنوات	فنادق ثابته			فنادق عائمه			قرى سياحية			الاجمالى		
	وحدات	غرف	أسره	وحدات	غرف	أسره	وحدات	غرف	أسره	وحدات	غرف	أسره
١٩٩٣	٤٣٧	٣٧٤٩٤	٧٣٢٩٠	١٨٨	٩٧٦٣	١٩٦٤٧	٦٧	٩٥٥١	١٩٣١١	٩٩٢	٥٨٧٥٥	١١٦٥٣١
١٩٩٤	٤٣٥	٤٩٤٤١	٧٧٥٢٩	٢٠٥	١٠٣٣٩	٢٠٥٣٣	٧٨	١٠٣٤٩	٢٠٩٣٤	٧١٨	٦١٠٦٨	١٢٠٨٥٤
١٩٩٥	٤٦١	٤٠٣٨٠	٧٩٣٨٧	٢٠٦	١٠٥٣٢	٢٠٩١٨	٨٥	١٠٦٩٨	٢١٧٧٧	٧٥٢	٦٤٩٥٨	١٢٨٩٥٧
١٩٩٦	٤٨٦	٤٣٧٢٨	٨٦٢٦٢	٢١٥	١١١٨٤	٢٢٣٣٣	٨٨	١١٧١٤	٢٣٦١٣	٧٢٩	٧٠٤٧١	١٤٠٧٤١
١٩٩٧	٥١٥	٤٧٥٧٣	٩٤٧٩٥	٢١٤	١١٣٢٢	٢٢٦٢٥	١٠٠	١٤١١٨	٢٨٤٦١	٨٢٩	٧٥٦٧٩	١٥٠٩٨٦
١٩٩٨	٥٢٢	٥٠٢٣٩	٩٩٩٠٠	٢٢٤	١١٦٧٣	٢٣٢١٦	١٢٣	١٨٧٨٤	٣٨٩٩٢	٨٦٩	٨٢٩٢٥	١٦٦٨١٧
١٩٩٩*	٥٣٤	٥٢٤٦٨	١٠٤٧٠٠	٢٣٠	١١٦٨٥	٢٣٣٢٠	١٣٠	١٨٨٩٥	٣٩٠٩٥	٨٩٤	٨٣٠٤٨	١٦٧١١٥

(*) : أرقام تقديرية

المصدر : - وزارة السياحة - النشرات السنوية للسياحة فى أرقام - أعداد مختلفة



٣- عرض البيانات فى صورة جدولية وبيانية : -

يعتبر عرض البيانات جدوليا وبيانيا هو الأكثر شيوعا خاصته فى البحوث والدراسات العلمية ، وجدير بالإيضاح أن العرض البيانى للبيانات لابد وان يسبق عرض جدولى سبق تصميمه .

أولا : عرض البيانات جدوليا وبيانيا فى حالة المجتمع الإحصائي الصغير : -

يتم اللجوء إلى استخدام هذا العرض للبيانات إذا كانت مفـــردات المجتمع الإحصائي محل الدراسة ذات عدد قليل ، ويرغـــب البـــاحث أن يعرضها فى جدول بسيط بدلا من تناولها فى سياق الكلام ، وأيضا ليتمكن من عرضها بيانيا فيسهل تبيان حقائق الأرقام واستقراء اتجاهاتها العامـــة بمجرد النظر ويتم عرض الجدول البسيط بيانيا فى عدة أشكال أهمها : -

١. الأعمدة البيانية .

٢. الدائــــــرة .

٣. الخط البيانى .

٤. خريطة الشريط .

١. الجدول البسيط وعرضة بيانيا فى شكل أعمدة : -

الأعمدة البيانية Bar charts هى عبارة عـــن أعمـــدة راســية تتناسب ارتفاعاتها مع قيم الظاهرة بشرط تساوى قواعد الأعمدة . ويتـــم فى هذه الطريقة تمثيل الصفة المميزة على المحور الأفقي ، ويتم تمثيـــل قيم الظاهرة على المحور الأسى .



مثال ١

الجدول التالى يمثل عدد الطلاب المقيدين باحدى الجامعات خلال الفترة (٩٥ – ١٩٩٩) .

(العدد بالألف)

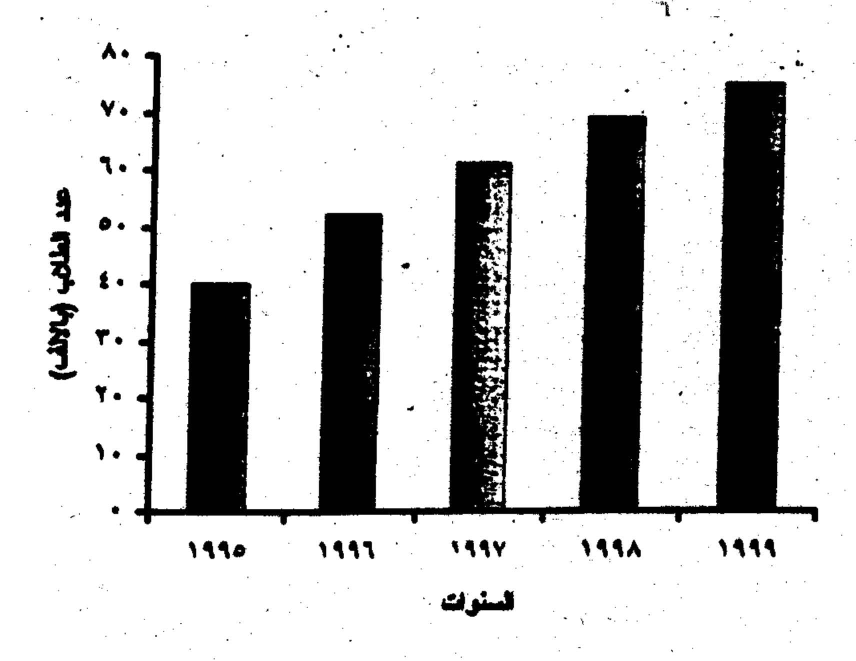
السنة الدراسية	١٩٩٥	٩٦	٩٧	٩٨	١٩٩٩
عدد الطلاب	٤٠	٥٢	٦١	٦٩	٧٥

المطلوب : -

عرض هذه البيانات في شكل أعمدة بيانية .

الحـل

تجدر الإشارة عند الرسم البيـــانى أن يســتخدم مقيـــاس رســم مناسب على المحوريـــن المتعـــامدين ، ومعنـــى كلمـــة مناســب أن ورقـــة الرســم البيـــانى تســتوعب المحـــور تقريبـــا وان المحـــور يستوعب اكبر رقم فى البيانات الممثلـــة عليـــه .



مثال ٢

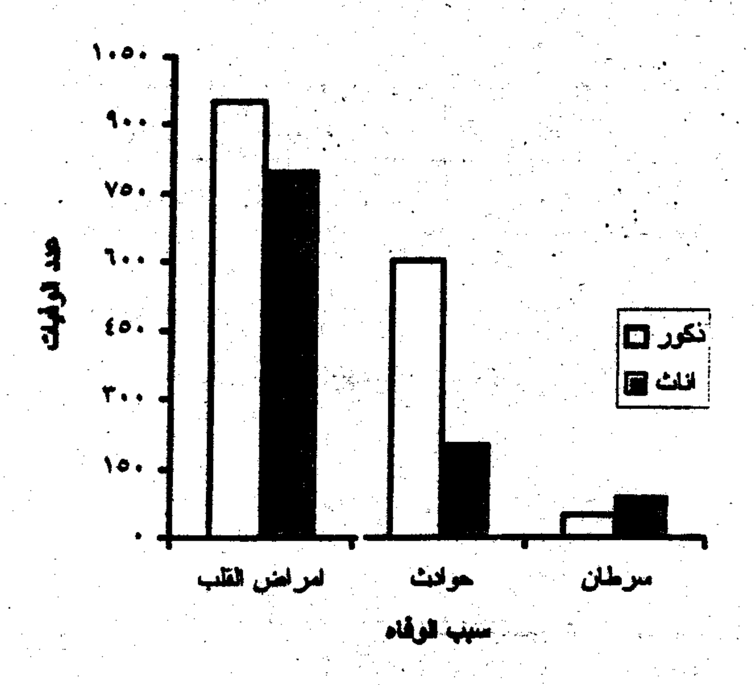
الجدول التالى يمثل عدد الوفيات حسب سببها (أمراض قلب ، حوادث ، سرطان) وذلك فى احدى المدن عام ١٩٩٥

سبب الوفاة	عدد الوفيات	
	ذكور	اناث
امراض القلب	٩٥٠	٨٠٠
حوادث	٦٠٠	٢٠٠
سرطان	٥٠	٩٠

المطلوب : -

عرض هذا الجدول بيانيا فى شكل أعمدة بيانية .

الحــــل



٢- الجدول البسيط وعرضة بيانيا فى شكل دائرة : Pie chart

تستخدم هذه الطريقة إذا كانت البيانات المعطاه عبارة عن مجموع عام مقسم إلى اجزاء فرعية ، وفى هذه الطريقة يتم تمثيل المجموع الكلى للبيانات بالمساحة الكلية للدائرة ، ثم يتم تقسيم الدائرة إلى قطاعات تتناسب فى مساحاتها مع المقادير الجزئية المكونة للمجموع الكلى وذلك من خلال تحويل المقادير الجزئية إلى نسب مئوية وضرب كل نسبة مئويـــة فـــى ٣٦٠ فيتم الحصول على الدرجات الستينية التى تستخدم فى الرسم البيانى للزاوية المركزية المعبرة عن القطاع الدائرى .

مثال- ١

الجدول التالى يبين عدد طلاب السياحة والفنادق فى عام ما موزعة على الأقسام الثلاثة : -

بيان	قسم السياحة	قسم الفنادق	قسم الارشاد	المجموع الكلى
عدد الطلاب	٢٥٠	٢٠٠	٥٠	٥٠٠

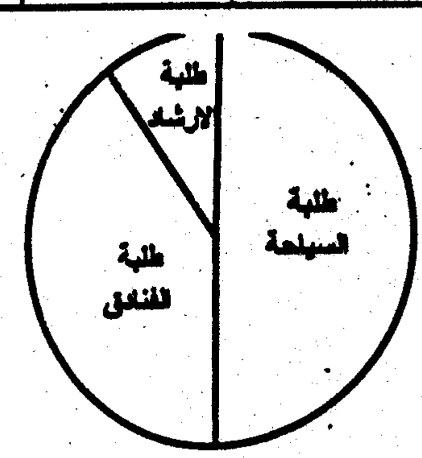
المطلوب :-

عرض هذه البيانات فى شكل دائرة .

الحــــل

تكوين الجدول الإحصائي اللازم

بيان	عدد الطلاب	النسبة المئوية %	الزاوية المركزية للقطاع الدائرى
طلبة السياحة	٢٥٠	٥٠	°١٨٠
طلبة الفنادق	٢٠٠	٤٠	°١٤٤
طلبة الارشاد	٥٠	١٠	°٣٦
المجموع	٥٠٠	١٠٠	°٣٦٠





مثال ٢

الجدول التالى يبين عدد طلاب السياحة والفنادق فى عام ما موزعة علـــى الأقسام الثلاثة ووفقا لنوع الطلبة بنين وبنات .

بيـــان	قسم السياحة			قسم الفنادق			قسم الارشاد		
	بنين	بنات	إجمالي	بنين	بنات	إجمالي	بنين	بنات	إجمالي
عدد الطلاب	١٦٠	٩٠	٢٥٠	٦٠	١٤٠	٢٠٠	٢٥	٢٥	٥٠

المطلوب : -

عرض هذه البيانات فى شكل دائرة .

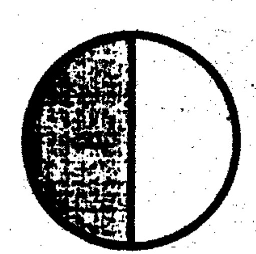
الحـــل

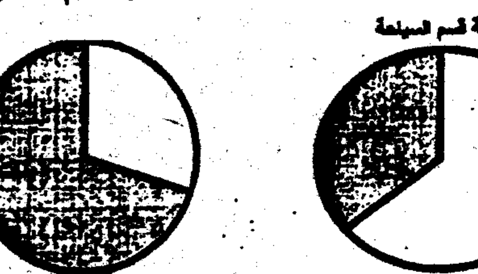
لعرض هذه البيانات يستلزم رسم ٣ دوائر متساوية .

تكوين الجدول الإحصائي اللازم : -

بيـــان	قسم السياحة			قسم الفنادق			قسم الارشاد		
	بنين	بنات	إجمالي	بنين	بنات	إجمالي	بنين	بنات	إجمالي
عدد الطلاب	١٦٠	٩٠	٢٥٠	٦٠	١٤٠	٢٠٠	٢٥	٢٥	٥٠
النسبة المئوية	٦٤	٣٦	١٠٠	٣٠	٧٠	١٠٠	٥٠	٥٠	١٠٠
الزاوية المركزية للقطاع الدائرى	٢٣٠	١٣٠	٣٦٠	١٠٨	٢٥٢	٣٦٠	١٨٠	١٨٠	٣٦٠

طلبة قسم الارشاد طلبه قسم الفنادق طلبة قسم السياحة





مثال ٣

الجدول التالى يبين توزيع جملة الإنفاق الحكومى على الخدمات الحكومية فى إحدى الدول خلال عامى ١٩٩٥ ، ٢٠٠٠ .

بيان	الأجور	المصروفات الإدارية	الاستثمارات	التحويلات	الجملة
١٩٩٥	١٤٠	٧٠	٨٠	٤٦	٣٦٦
٢٠٠٠	١٧٠	٨٠	٨٦	٦٤	٤٠٠

المطلوب : -

عرض هذه البيانات فى شكل دائرة .

يترك كتدريب .

٢.الجدول البسيط وعرضة بيانيا فى شكل خط بيانى : Line chart

يستخدم الخط البيانى إذا كانت البيانات تعبر عن سلوك ظاهرة ما مع الزمن، ويتم تمثيل الزمن كمتغير مستقل على المحور الأفقي ، بينما يتم تمثيل قيم الظاهرة كمتغير تابع على المحور الرأسى .

مثال

الجدول التالى يبين عدد القرى السياحية فى ج.م.ع خلال الفترة (١٩٩٣-١٩٩٩) .

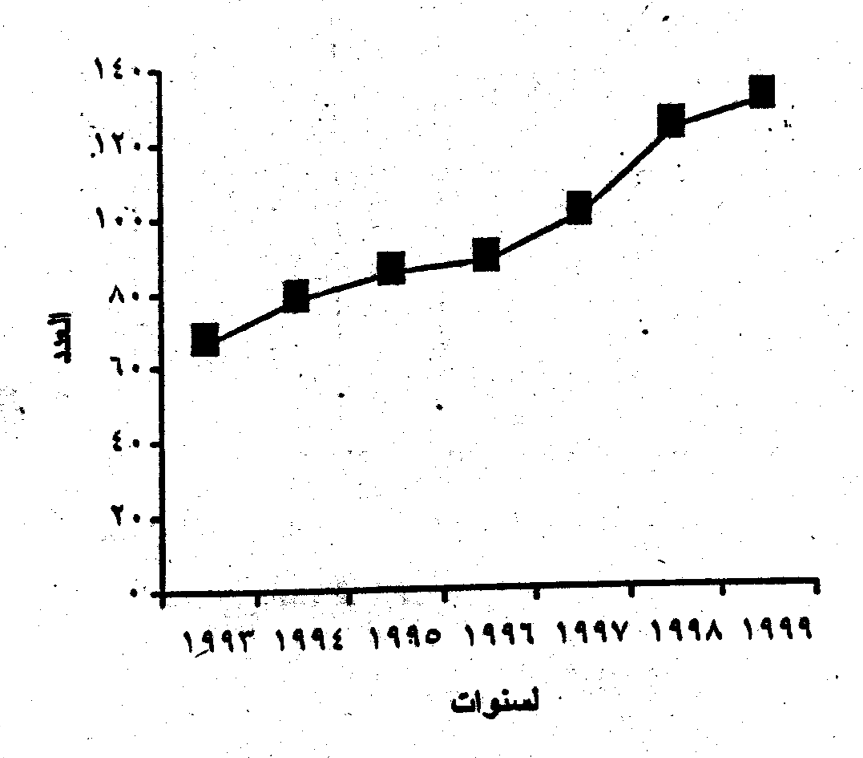
السنوات / عدد القرى السياحية	١٩٩٣	١٩٩٤	١٩٩٥	١٩٩٦	١٩٩٧	١٩٩٨	١٩٩٩
عدد القرى السياحية	٦٧	٧٨	٨٥	٨٨	١٠٠	١٢٣	١٣٠



المطلوب :-

عرض هذه البيانات فى شكل خط بيانى .





وهنا يلاحظ ضرورة استخدام مقياس الرسم المناسب .

٤. الجدول البسيط وعرضة بيانيا فى شكل خريطة الشريط Band chart

تستخدم خريطة الشريط إذا كانت البيانات تعبر عن سلوك ظاهرتين من نفس النوع مع الزمن ، وعلى ذلك فخريطة الشريط توضح تطور الظاهرتين معا مع الزمن وأيضا توضح تطور الفرق بينهما .



مثال

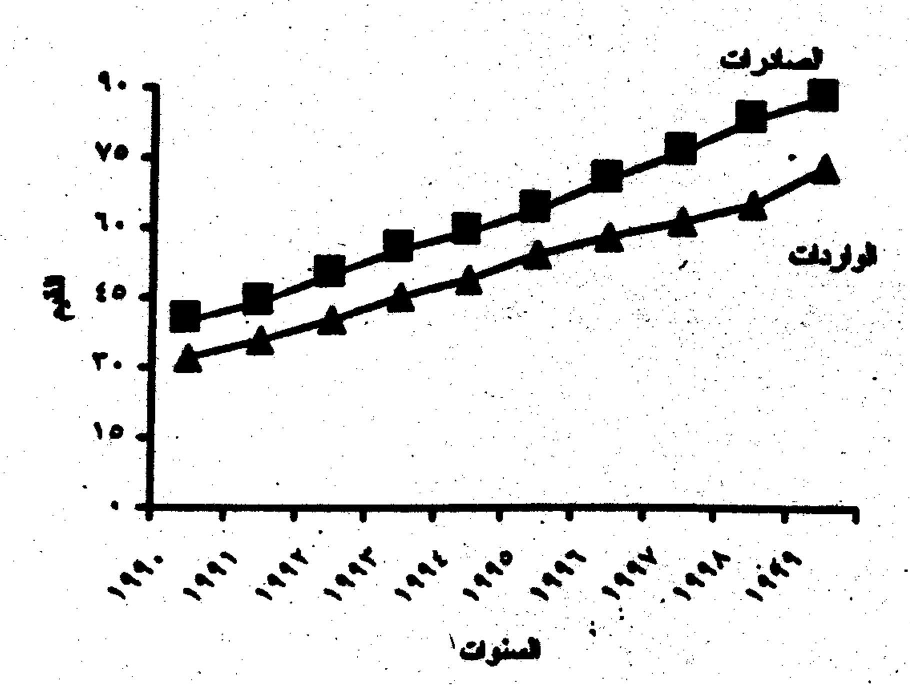
الجدول التالى يبين تطور قيمة الصادرات والواردات لدولة ما خلال الفترة (١٩٩٠-١٩٩٩) بالمليون دولار .

السنوات / بيان	١٩٩٠	١٩٩١	١٩٩٢	١٩٩٣	١٩٩٤	١٩٩٥	١٩٩٦	١٩٩٧	١٩٩٨	١٩٩٩
الصادرات	٤٠	٤٤	٥٠	٥٥	٦٤	٦٤	٧٠	٧٦	٨٣	٨٨
الواردات	٣٢	٣٦	٤٠	٤٥	٤٩	٥٤	٥٨	٦١	٦٥	٧٣
صافى الميزان	+٨	+٨	+١٠	+١٠	+١٠	+١٠	+١٢	+١٥	+١٨	+١٥

المطلوب : -

عرض هذه البيانات بيانيا .

الحـــل





ثانيا : عرض البيانات جدوليا وبيانيا فى حالة المجتمع الإحصائي الكبير : -

يتم اللجوء إلى استخدام هذا العرض للبيانات إذا كانت مفردات المجتمع الإحصائي ذات عدد كبير مما يصعب معه دراستها إحصائيا إلا من خلال عرض جدولى وبيانى مناسب ،ولكى نتحقق كلمة مناسب هذه يجب معرفة الاتى : -

١ تكوين جدول التوزيع التكرارى ، وعرضه بيانيا فى شكل المدرج التكرارى والمضلع التكرارى والمنحنى التكرارى

٢ تكوين جدول التوزيع التكرارى التجميعى الصاعد والهابط ، وعرضه بيانيا فى شكل المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط .

٣ تكوين جدول التوزيع التكرارى النسبي والمئوى ، وعرضه بيانيا فى شكل المنحنى النسبي والمنحنى المئوى .

وبهذا يمكن التعرف على خصائص المجتمع الإحصائي الكبير ، وفيما يلى التناول للنقاط الثلاث السابقة بالشرح والأمثلة .

١ تكوين جدول التوزيع التكرارى ، وعرضه بيانيا فى شكل المدرج التكرارى والمضلع التكرارى والمنحنى التكرارى

البيانات التالية تمثل الإنفاق اليومى لعدد ٢٠٠ سائح فى ج.م.ع بالدولار : -

٤	٩	٩	١١	٢١	٩	٥	٧	١٤	٣٠
٧	٢٢	٢٣	٢٤	٤٠	١١	٢٧	١٣	٨	١١
١٦	٢١	١٣	١٥	١٦	١٧	١٤	١٥	١٣	٨



١٧	١٢	١٤	١٨	١٩	١٩	١٦	١٢	٨	١٣
١٨	٢٠	١٩	١٢	٢٠	١٨	١٧	١٣	١٢	١٦
١٩	١٧	٢٢	٢١	١٩	٢١	٢٠	١٨	١٦	١٧
١٧	١٨	١٩	٢٢	١٢	٢٢	٢١	٢٢	١٧	١٨
٢٠	١٩	٢٢	٢٣	٢٣	٢١	٢٢	١٨	٢٠	١٩
١٨	٢٢	٢١	٢٤	٢٤	٢٣	١٨	٢٠	١٨	٢١
٢٢	٢١	٢٠	٢٤	٢٣	٢٢	٢٠	٢٤	٢١	٢٠
٢٤	٢٣	٢٤	٢٣	٢٤	٢٠	٢٥	٢٣	٢٢	٢٣
٢٤	٢٢	٢٣	٢٢	٢٥	٢١	٢٢	٢٠	٢٥	٢١
٢٥	٢٣	٢٢	٢٥	٣٨	٢٤	٢٣	٢٥	٢٤	٢٣
٢٥	٢٤	٣٥	٢٤	٢٥	٣٥	٢٤	٢٣	٢٤	٢٥
٣٦	٣٣	٣٢	٢٥	٣٢	٣٤	٣٠	٢٤	٣٣	٢٤
٣١	٢٨	٢٧	٢٦	٣١	٣٢	٣١	٣٣	٣١	٣٣
٢٧	٢٦	٢٧	٢٨	٢٦	٢٨	٢٦	٢٧	٣٣	٢٦
٢٦	٢٩	٢٩	٢٦	٢٩	٢٦	٢٨	٢٨	٢٧	٢٧
٢٩	٣٠	٢٦	٣٠	٢٧	٢٩	٢٦	٢٨	٢٦	٢٨
٢٨	٣٠	٢٩	٢٧	٢٨	٢٧	٢٩	٢٧	٢٧	٣٤

المطلوب : -

أولا : - عرض و تلخيص البيانات فى جدول توزيع تكرارى

ثانيا : - عرض هذا الجدول بيانيا فى شــكل مــدرج تكــرارى ، مضلع تكرارى ، منحنى تكرارى .



الحـــل

تكوين جدول التوزيع التكرارى :

لعرض البيانات السابقة فى جدول توزيع تكرارى نتبع الاتى : -

١- يتم اختيار عدد معين من الفئات ، وهناك اتفاق بين الباحثين على أن عدد الفئات يتراوح بين ٨ فئات إلى ١٢ فئة ، ولا توجد قاعدة فى عملية الاختيار هذه اللهم شرط أن يكون عدد الفئات مناسب ، ومعنى كلمة مناسب ألا يكون عدد الفئات قليل حتى لا يضيع الكثير من التفاصيل اللازم معرفتها عن معالم التوزيع ، والا يكون عدد الفئات كثير فتضيع الحكمة من عمل جدول التوزيع التكرارى وهى التلخيص لمعرفة معالم التوزيع .

٢- إيجاد المدى للبيانات وهو الفرق بين اكبر قيمة واصغر قيمة للبيانات .

٣- يحسب طول الفئة * وذلك بقسمة المدى على عدد الفئات المختارة .

٤- يتم تحديد الفئة الأولى كما يلى : -

<u>الحد الادنى للفئة الأولى :</u> وهو اصغر قيمة فى البيانات الموجودة .

<u>الحد الاعلى للفئة الأولى :</u> وهو اصغر قيمة فى البيانات الموجودة مضافا إليه طول الفئة المحسوب

---

* يسمى طول الفئة بمدى الفئة أو اتساع الفئة



٥- تحدد الفئات بعد ذلك كما يلى : -

يضاف طول الفئة إلى الحد الاعلى للفئة الأولى فينتج الحد الادنى للفئة الثانية ، وهكذا .

٦- يتم عمل الجدول الإحصائي اللازم كما يلى : -

الفئات	العلامات	التكرار
(٠ - ٥) (٥ - ١٠) (١٠ - ١٥)		
المجموع		

ويلاحظ أن الفئات قد تكتب بطرق أخرى غير الموجودة بالجدول كأن تكتب على الصورة (٠ -) ، (٥ -) ، (١٠ -) أو على الصورة (- ٥) ،(- ١٠) ، (- ١٥) .

٧- الرصد للبيانات فى الجدول وذلك بالتعبير عن كل قيمة فى البيانات بعلامة مائلة أمام الفئة المناسبة لها ، وكل ٤ علامات يتم قفلها بعلامة خاصة معكوسة لتكوين حزمة لتعبر عن ٥ مفردات .

وعند رصد القيمة ١٠ فلا يتم رصدها فى الفئة الثانية وانما يتم رصدها فى الفئة الثالثة والسبب أن الفئة الثانية هى من ٥ إلى اقل من ١٠ .

بعد ذلك يتم جمع العلامات الموجودة امام كل فئة وترصد فى عمود التكرار فى صورة رقمية .



جدول رقم (   ) : توزيع الإنفاق اليومى لعدد ٢٠٠ سائح في ج.م.ع (بالدولار)

الفئات	العلامات	التكرار
(٠ –٥)	//	٢
(٥ –١٠)	/// ////	٨
(١٠–١٥)	// //// //// ////	١٧
(١٥–٢٠)	/ //// //// //// //// //// //// //// //// ////	٤١
(٢٠–٢٥)	//// //// //////// //// //// //// //// //// //// //// //// ////////	٦٩
(٢٥–٣٠)	//// //// //// //// //// //// //// //// ////	٤٤
(٣٠–٣٥)	/ //// //// ////	١٦
(٣٥–٤٠)	///	٣
المجموع		٢٠٠

وقد تم تكوين الجدول السابق على أساس أن اصغر رقم فى البيانات هو ٤ وان اكبر رقم فيها هو ٤٠ ، وعلية فالمدى يساوى ٣٦ ، وبالتالى فطول الفئة هو ناتج قسمة ٣٦ على ٨ وذلك لاختيار عدد الفئات ٨ ، وعلى ذلك فطول الفئة هو ٥ أي تقريب ناتج القسمة ٤,٥ إلى رقم صحيح .

ويلاحظ فى الجدول أن مجموع التكرارات والذى يساوى ٢٠٠ هو عدد البيانات المراد عرضها فى جدول التوزيع التكرارى .

٨– بعد عمل جدول التوزيع التكرارى للبيانات المعبرة عن الإنفاق اليومى لعدد ٢٠٠سائح اصبح انفاق كل سائح على حده ليس له وجود ، و إنما اصبح واضحا الانفاق لمجموعة من السائحين فى فئة معينة ، فمثلا نقول يوجد ١٧ سائح



يتراوح إنفاقهم اليومى بين (١٠ – ١٥) دولار . إلا انه يوجد وصف افضل كأن نقول يوجد ١٧ سائح يبلغ متوسط انفاقهم اليومى ١٢,٥ دولار ، والرقم الأخير هو ناتج قسمة مجموع (الحدين الأدنى والأعلى للفئة) على ٢ ، وهذا الناتج يسمى بمركز الفئة .

.. مركز الفئة = (الحد الادنى + الحد الاعلى للفئة) / ٢

والجدول السابق يسمى جدول مقفول من طرفيه حيث الطرف الأول وهو الفئة الدنيا معروف بدايتها ونهايتها ، وكذلك الطرف الآخر وهو الفئة العليا معروف بدايتها ونهايتها . أما إذا كان جدول التوزيع التكرارى طرفة الأول أي الفئة الدنيا مكتوبة على الصورة (– ٥) أو على الصورة (٥ –) وكذلك طرفة الأخر أي الفئة العليا مكتوبة على نفس الصورة فيقال أن هذا التوزيع مفتوح سواء من طرف واحد أو من طرفية معا .

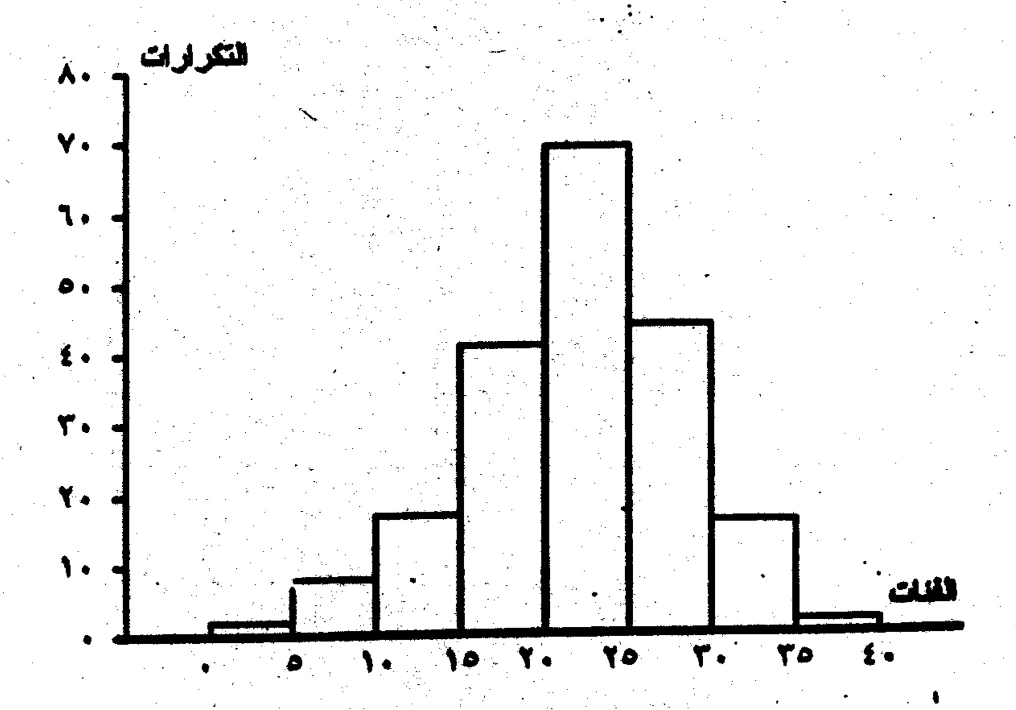
المدرج التكرارى : Histogram

المدرج التكرارى هو العلاقة البيانية بين حدود الفئات وتكراراتها بحيث تمثل حدود الفئات على المحور الأفقي وتكراراتها على المحور الرأسى .

ولعرض جدول التوزيع التكرارى السابق فى شكل مدرج تكرارى يتم تقسيم المحور الأفقي إلى تقسيمات تساوى عدد الفئات ، وان التقسيم عبارة عن شرطة تعبر عن طول الفئة ، وإذا كانت الفئات اطوالها متساوية فان التقسيمات تكون متساوية ، ثم يتم توقيع قيم تكرارات المحور الصادى بشرطة توازى وتساوى شرطة طول الفئة المناظر ، ثم تـ تكمل



مستطيلات هاتين الشرطتين المتوازيتين فينشأ المدرج التكراري كما يلي:-



ملاحظات :-

١. يتم ترك مسافة ٢ سم من نقطة الأصل على المحور الأفقي .

٢. اختلاف الرسم البياني للتوزيع التكراري عن الرسم البياني العادي ، فالأخير عبارة عن نقط ($س_1$ ، $ص_1$) ، ($س_2$ ، $ص_2$) ، .... وأما الأول فهو عبارة عن مستطيلات عروضها متساوية إلا أن أطوالها مختلفة لتنشأ مساحات مستطيلات تتناسب وتكرارات فئات التوزيع .

٣. أن مجموع مساحات المستطيلات تساوي مجموع التكرارات .

٤. إذا كانت الفئات غير متساوية الطول فان الرسم البيانى النـــاتج سيكون مضللاً لذلك يجب إدخال تعديلات مناسبة فى طريقـــة الرسم البيانى لتلافى هذا التضليل .

مثال

الجدول التالى هو جدول توزيع تكرارى

الفئات	(١٠-٢٠)	(٢٠ - ٣٠)	(٣٠-٤٠)	(٤٠-٦٠)	(٦٠-١٠٠)	(١٠٠-١٥٠)
التكرارات	٥٠	١٠٠	٢٠٠	٣٠٠	٤٠٠	٢٥٠

المطلوب : -

عرض هذه البيانات وتدوين الملاحظات .

الحـــل

التكرارات

٤٠٠
٣٠٠
٢٠٠
١٠٠
٠

الفئات

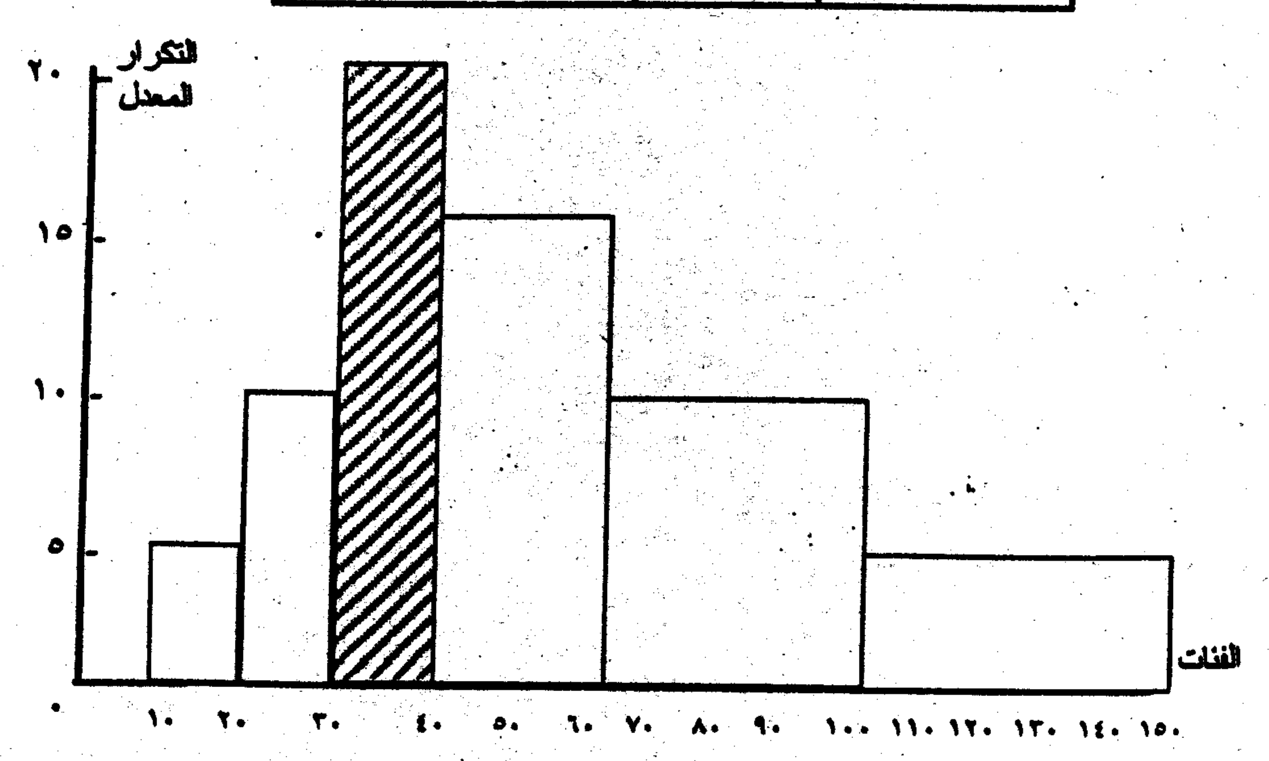
١٠ ٢٠ ٣٠ ٤٠ ٥٠ ٦٠ ٧٠ ٨٠ ٩٠ ١٠٠ ١١٠ ١٢٠ ١٣٠ ١٤٠ ١٥٠



وهنا يلاحظ أن الفئة (٦٠ - ١٠٠) بها اكثر التكرارات أي ٤٠٠ مفرده وفى هذا تضليل حيث أن طول هذه الفئة اكبر من اطوال فئات أخرى ، ولذلك وجب تلافى هذا التضليل كما يلى : -

الحل

الفئات	التكرارات	التكرارات المعدلة *
(١٠ - ٢٠)	٥٠	٥
(٢٠ - ٣٠)	١٠٠	١٠
(٣٠ - ٤٠)	٢٠٠	٢٠
(٤٠ - ٦٠)	٣٠٠	١٥
(٦٠ - ١٠٠)	٤٠٠	١٠
(١٠٠ - ١٥٠)	٢٥٠	٥



* التكرار المعدل = $\frac{\text{تكرار الفئة}}{\text{طول الفئة}}$ ، وبهذا يوجد أساس واحد لكل الفئات

وهنا يلاحظ أن الفئة (٣٠ – ٤٠) هى التى بها اكثر التكرارات .

المضلع التكرارى Frequency

المضلع التكرارى هو العلاقـة البيانيـة بيـن مراكـز الفئـات وتكراراتها بحيث تمثل مراكز الفئات على المحور الأفقي وتكراراتها على المحور الرأسى .

والرسم البيانى فى هذه الحالة هـو الرسـم البيـانى العـدى و احداثياتها الرأسيه هى تكرارات هذه الفئات ، ثم يتم توصيل هـذه النقـط بخطوط مستقيمة فنحصل على المضلع التكرارى .

مثال

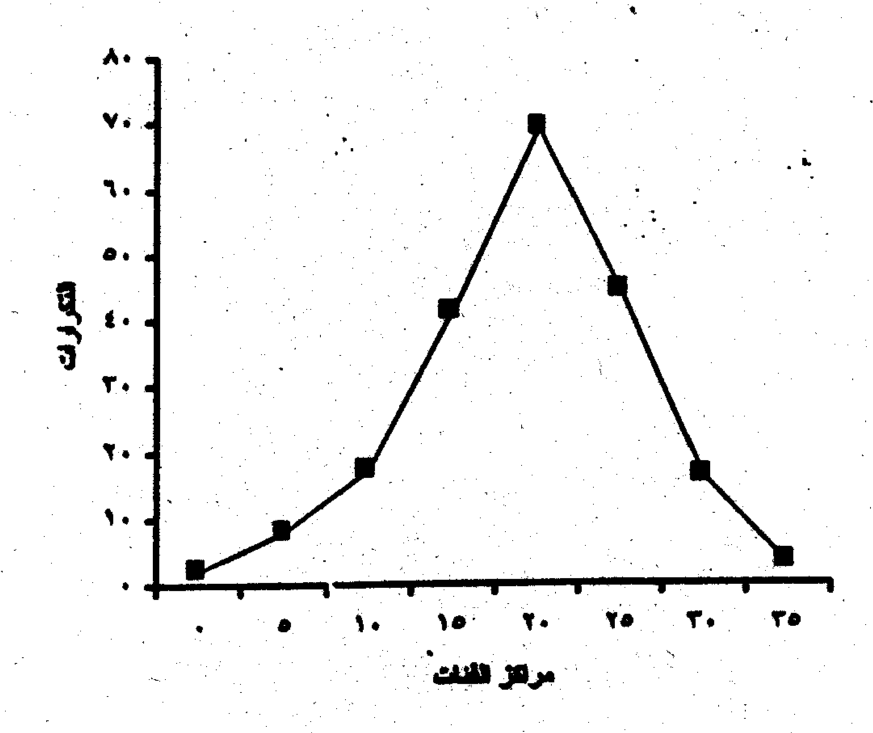
المطلوب العرض البيانى لجـدول التوزيـع التكـرارى ذو الفئـات المتساوية السابقه فى شكل مضلع تكرارى

تكوين الجدول الإحصائي اللازم : -

الفئات	التكرارات	مركز الفئة
(٠ – ٥)	٢	٢,٥
(٥ – ١٠)	٨	٧,٥
(١٠–١٥)	١٧	١٢,٥
(١٥ – ٢٠)	٤١	١٧,٥
(٢٠–٢٥)	٦٩	٢٢,٥
(٢٥–٣٠)	٤٤	٢٧,٥
(٣٠–٣٥)	١٦	٣٢,٥
(٣٥–٤٠)	٣	٣٧,٥

العرض البيانى للمضلع التكرارى :-





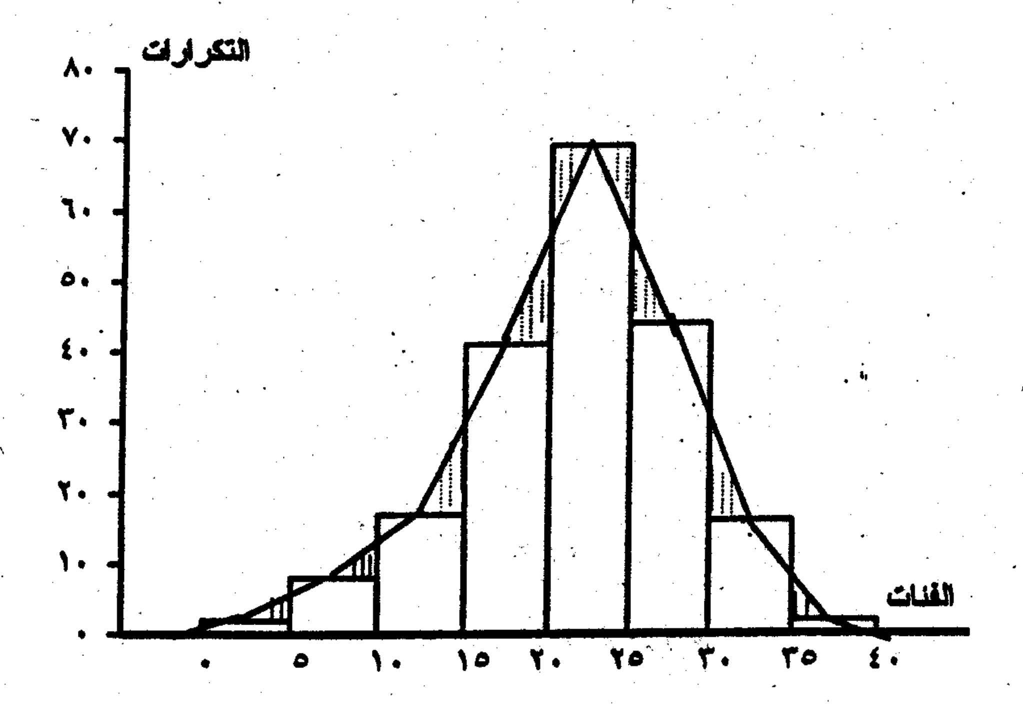
ملاحظات : -

١. أن المساحة تحت المضلع التكراري تساوى المساحة تحت المدرج التكرارى وبالتالى تساوى مجموع التكرارات .

الاثبات

فى المدرج التكرارى السابق يتم تنصيف الاضلاع العليا للمستطيلات فى نقط ، وبالطبع هذه النقط احداثياتها الأفقية هى مراكز الفئات واحداثياتها الراسية هي التكرارات المناظرة ، وانه بتوصيل هذه النقط بخطوط مستقيمة نحصل على المضلع التكرارى كما يلى : -





ويتضح من الرسم أن مجموع مساحات ΔΔΔ الخارجة عن المدرج التكرارى هى فى نفس الوقت داخله ضمن مساحة المضلع التكرارى (المثلثات المنقطة) ، وان مجموع مساحات ΔΔΔ الداخلة ضمن المدرج التكرارى هى نفس الوقت خارجة عن مساحة المضلع التكرارى ، ولما كان باقى مساحات المستطيلات مشتركة بين المدرج التكرارى والمضلع التكرارى ، إذا مساحة المدرج التكرارى تساوى المساحة تحت المضلع التكرارى .

٢. جدير بالذكر أن المضلع التكرارى يمكن استخدامه عند إجراء مقارنات بين توزيعين تكراريين مختلفين فى حين يصعب إجواء مثل هذه المقارنة باستخدام المدرجات التكرارية ، كما يعتبر المضلع التكرارى هو الخطوة الأولى فى رسم المنحنى التكرارى

المنحنى التكرارى **Frequency curve**

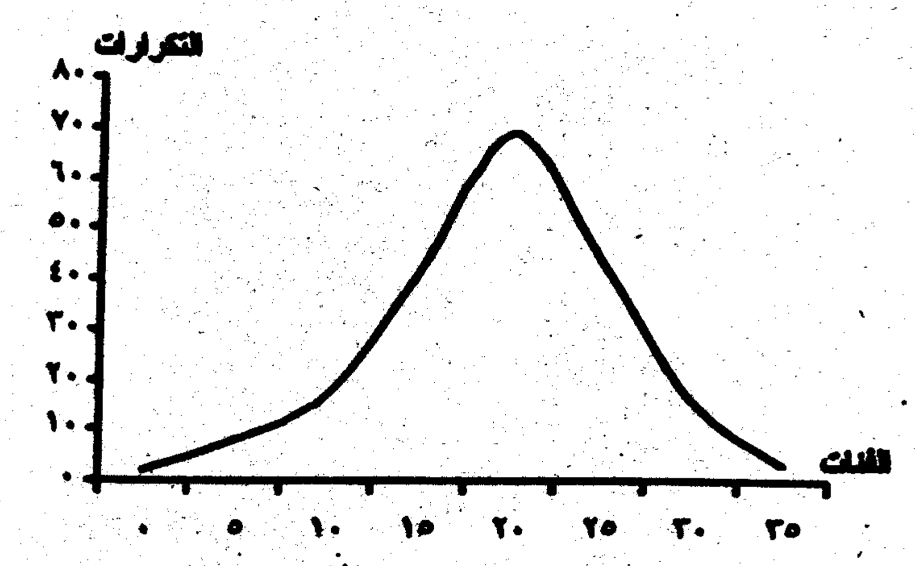
لقد سبق القول أن المنحنى التكرارى يتم رسمه من المضلع التكرارى ، لكن كيف يتأتى ذلك .

مثال

المطلوب العرض البياني لجدول التوزيع التكرارى ذو الفئات المتساوية السابق في شكل منحنى تكرارى .

الحل

يتم رسم المضلع التكرارى كما سبق ، ثم يتم التعامل مع الخطوط المستقيمة لهذا المضلع بالتمهيد باليد لتصبح خطا مستقيما يمر بأكبر عدد ممكن من النقط وعلى أن يمر خلال باقي النقط بالتوازن فيما بينها أي نقطة فوق ونقطة تحت كما يلى : -



وجدير بالذكر أن المساحة تحت المنحنى التكرارى هى أيضا تساوى المساحة تحت المضلع التكرارى وبالتالى تساوى مجموع

التكرارات ، إلا انه لاثبات ذلك فان الأمر يتطلب معطيات أخرى عن المنحنى التكرارى حتى يمكن استخدام حساب التكامل فى إيجاد المساحة المحصورة بين منحنى وضلع ، هذا ما سيتم تناوله بأذن الله فى مواضع إحصائية متقدمة .

٢ تكوين جدول التوزيع التكرارى التجميعى الصاعد والهابط ، وعرضه بيانيا فى شكل المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط .

من الأهمية بمكان الإيضاح بان جدول التوزيع التكرارى التجميعى الصاعد والهابط يشتقان من جدول التوزيع التكرارى الأصلي ، ونلجأ إليهما عندما نريد معرفة عدد المفردات التى تقل عن قيمة مفردة معينة ، أو عدد المفردات التى تزيد عن قيمة مفرده معينة ، أو عدد المفردات التي تنحصر بين قيمتى مفردتين معينتين .

مثال

الجدول التالى هو جدول توزيع تكرارى .

الفئات	(٠-٥)	(٥-١٠)	(١٠-١٥)	(١٥-٢٠)	(٢٠-٢٥)	(٢٥-٣٠)	(٣٠-٣٥)	(٣٥-٤٠)	المجموع
التكرارات	٢	٨	١٧	٤١	٦٩	٤٤	١٦	٣	٢٠٠

المطلوب : -

عرض بيانات هذا الجدول التكرارى فى شكل جدول توزيع تكرارى مجمع صاعد وهابط .



الحــــل

الفئات	التكرارات	الحدود العليا للفئات * وتباينات أصغر من	التكرار المتجمع الصاعد
		> ٠	صفر
(٠ - ٥)	٢	> ٥	٢
(٥ - ١٠)	٨	> ١٠	١٠
(١٠-١٥)	١٧	> ١٥	٢٧
(١٥ - ٢٠)	٤١	> ٢٠	٦٨
(٢٠-٢٥)	٦٩	> ٢٥	١٣٧
(٢٥-٣٠)	٤٤	> ٣٠	١٨١
(٣٠-٣٥)	١٦	> ٢٥	١٩٧
(٣٥-٤٠)	٣	> ٤٠	٢٠٠
المجموع	٢٠٠		

ملاحظـــــــات :-

١. تم تكوين الجدول التكرارى المتجمع الصاعد باستـــتخدام الحــدود العليا للفئات ثم جمع التكرارات بالتالي (الجمع الـــتراكمى) مــن بداية التوزيع .

٢. أن هذا التوزيع يعطى وصفا اكثر عن ســلوك المتغـير محـل الدراسة وهو انفاق عدد ٢٠٠ سائح حيث يمكــن التعــرف علــى المعلومات التالية :-

* هذا العمود بمثابة خط اعداد لمفردات المجتمع ككل .



- أن عدد السائحين الذين يقل انفاقهم عن ١٥ دولار هو ٢٧ سائح .
- أن عدد السائحين الذين يقل انفاقهم عن ٣٠ دولار هو ١٨١ سائح .
- أن عدد السائحين الذين ينحصر انفاقهم بين ٢٥ ، ٣٥ دولار هو ٦٠ سائح .

وهكذا .

تكوين جدول التوزيع التكرارى المتجمع الهابط : -

الفئات	التكرارات	الحدود الدنيا للفئات وتباينات اكبر من *	التكرار المتجمع الهابط
(٠ - ٥)	٢	< ٠	٢٠٠
(٥ - ١٠)	٨	< ٥	١٩٨
(١٠-١٥)	١٧	< ١٠	١٩٠
(١٥ - ٢٠)	٤١	< ١٥	١٧٣
(٢٠-٢٥)	٦٩	< ٢٠	١٣٢
(٢٥-٣٠)	٤٤	< ٢٥	٦٣
(٣٠-٣٥)	١٦	< ٣٠	١٩
(٣٥-٤٠)	٣	< ٢٥	٣
	٠	< ٤٠	صفر
المجموع	٢٠٠		

*هذا العمود بمثابة خط اعداد لمفردات المجتمع ككل .

ملاحظات :-

١.تم تكوين جدول التكرار المجتمع الهابط باستخدام الحــــدود الدنيـــا للفئات ثم جمع * التكرارات بالتتالى (الجمع التراكمى) من نهايـــة التوزيع .

٢.أن هذا التوزيع يعطى وصفا اكثر عن سلوك المتغير حيث يمكـــــن التعرف على المعلومات التالية : -

- أن عدد السائحين الذين يزيد انفاقهم عن ١٥ دولار هـــــو ١٧٠ سائح .
- أن عدد السائحين الذين يزيد انفاقهم عن ٣٠ دولار هـــــو ١٩ سائح .
- أن عدد السائحين الذين ينحصر انفاقـــهم بيـــن ٢٥ ، ٣٥ دولار هو ٦٠ سائح .

*يمكن الطرح المتتالى من بداية التوزيع

العرض البياني :



٣ تكوين جدول التوزيع التكرارى النسبي والمئوى وعرضهما بيانيا فى شكل المنحنى التكراى النسبي والمنحنى التكرارى المئوى .

يشتق هذين التوزيعين من التوزيع التكرارى الأصلي ونلجأ إليها عندما نريد معرفة الوزن النسبي لكل فئة على حده .

مثال

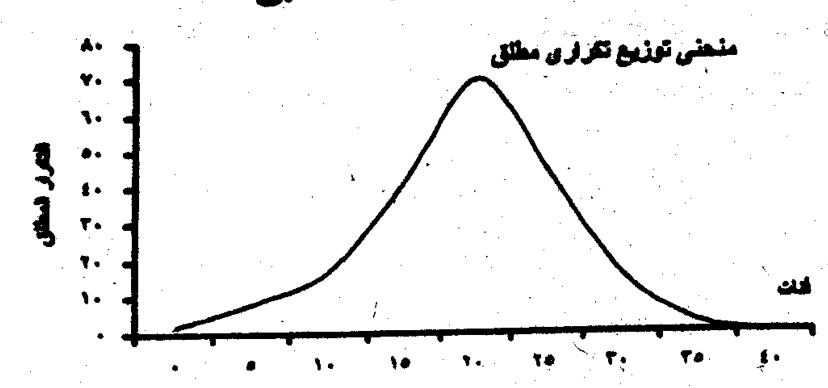
المطلوب عمل جدول توزيع تكرارى نسبي ومئوى لجدول التوزيع التكرارى الأصلي والخاص بانفاق عدد ٢٠٠ سائح .

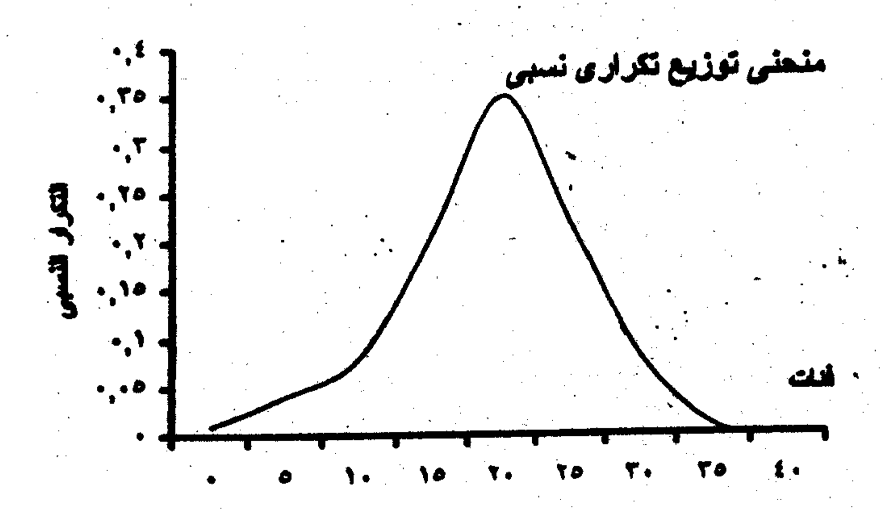
الحل

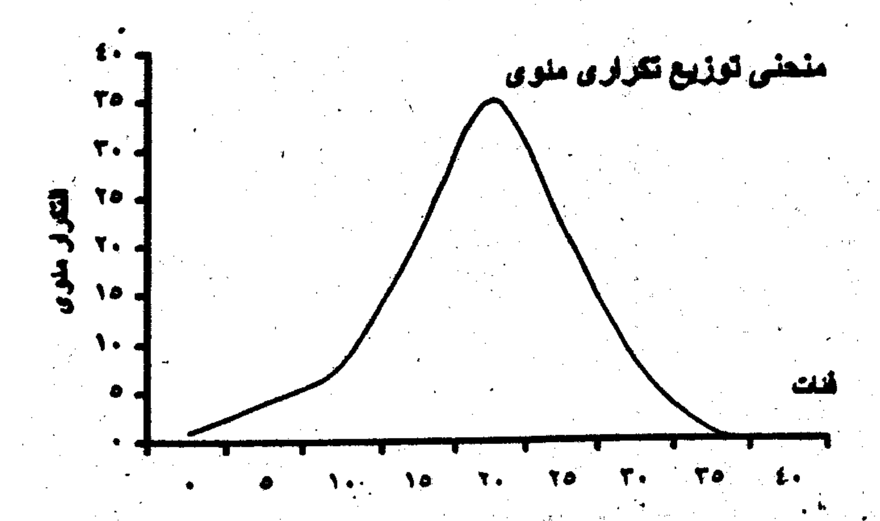
تكوين جدول التوزيع التكرارى النسبى والمئوى : -

فئات (الإنفاق)	التكرارات (عدد السائحين)	التكرار النسبي	التكرار المئوى (%)
(٠ - ٥)	٢	٠,٠١	١
(٥ - ١٠)	٨	٠,٠٤	٤
(١٠-١٥)	١٧	٠,٠٨	٨
(١٥ - ٢٠)	٤١	٠,٢١	٢١
(٢٠-٢٥)	٦٩	٠,٣٥	٣٥
(٢٥-٣٠)	٤٤	٠,٢٢	٢٢
(٣٠-٣٥)	١٦	٠,٠٨	٨
(٣٥-٤٠)	٣	٠,٠١	١
المجموع	٢٠٠	١	١٠٠

العرض البيانى للتوزيعات التكرارية بالجدول السابق : -









ملاحظـــــــات :-

أ - أن نمط التوزيع المطلق هو نمط التوزيع النسبي هو نمـــط التوزيـــع المئوى .

ب- أن الاختلاف بين الأشكال الثلاثة ينحصر فقط فى مقيـــاس الرســـم المستخدم على المحور الرأسى .

جـ - إذا كنا بصدد معرفة وصف الاختلاف فى البيانات كلـــها بنســـب مئوية ، فأننا نستخدم التوزيع التكرارى التجميعى الصاعد والـــهابط فى شكل نسب مئوية .

مثال

المطلوب عمل جدول توزيع تكرارى تجميعى صاعد وهابط فـــى شكل نسب مئوية

تكوين جدول التوزيع التكرارى المتجمع الصاعد فى شكل نسب مئوية:-

الفئات	التكرارات	الحدود العليا للفئات* وتباينات اصغر من	التكرار المتجمع الصاعد	التكرار المتجمع الصاعد المئوى %
		> ٠	صفر	٠
(٠ - ٥)	٢	> ٥	٢	١
(٥ - ١٠)	٨	> ١٠	١٠	٥
(١٠-١٥)	١٧	> ١٥	٢٧	١٣,٥
(١٥-٢٠)	٤١	> ٢٠	٦٨	٣٤
(٢٠-٢٥)	٦٩	> ٢٥	١٣٧	٦٨,٥
(٢٥-٣٠)	٤٤	> ٣٠	١٨١	٩٠,٥
(٣٠-٣٥)	١٦	> ٢٥	١٩٧	٩٨,٥
(٣٥-٤٠)	٣	> ٤٠	٢٠٠	١٠٠
المجموع	٢٠٠			

*هذا العمود بمثابة خط اعداد لمفردات المتجمع ككل .

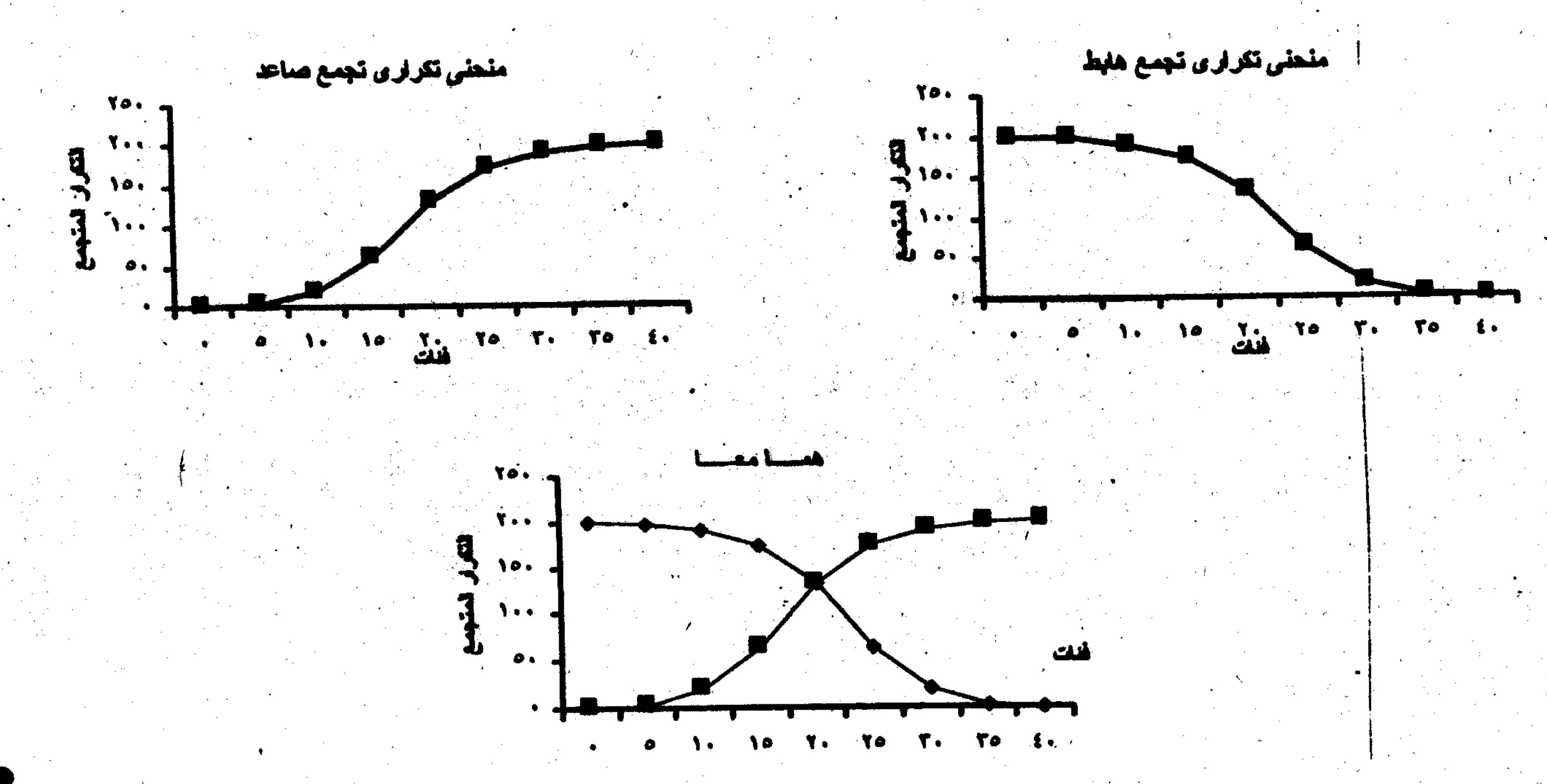


تكوين جدول التوزيع التكرارى المتجمع الهابط فى شكل نسب مئوية:-

الفئات	التكرارات	الحدود الدنيا للفئات* وتباينات اكبر من	التكرار المتجمع الهابط	التكرار المتجمع الهابط المئوى %
(٠ - ٥)	٢	< ٠	٢٠٠	١٠٠
(٥ – ١٠)	٨	< ٥	١٩٨	٩٩
(١٠-١٥)	١٧	< ١٠	١٩٠	٩٥
(١٥ – ٢٠)	٤١	< ١٥	١٧٣	٨٦,٥
(٢٠-٢٥)	٦٩	< ٢٠	١٣٢	٦٦
(٢٥-٣٠)	٤٤	< ٢٥	٦٣	٣١,٥
(٣٠-٣٥)	١٦	< ٣٠	١٩	٩,٥
(٣٥-٤٠)	٣	< ٢٥	٣	١,٥
		< ٤٠	صفر	٠
المجموع	٢٠٠			

* هذا العمود بمثابة خط اعداد لمفردات المتجمع ككل .

العرض البيانى لجدولى التوزيع التكرارى التجميعى الصاعد والهابط : -



ملاحظـات : -

١.من المهم جدا معاودة الاضاح بالاتى : -

أ) أن الرسم البيانى العادى هو علاقة بين متغيرين ، ومن ثـم تظهر هذه العلاقة جدوليا فـى شـكل نقـط (س ، ص) ، (س ، ص) ،... . وبيانيا فى شكل خط بيانى مسـتقيم أو منحنى .

ب) أن الرسم البيانى للتوزيع التكرارى هو علاقة بين حدود فئات المجتمع على المحور الأفقي والتكرارات المناظرة لها علـى المحور الرأسى ، ومن ثم تظهر هذه العلاقة جدوليا فى شكل (الفئة الأولى وتكرارها المناظر) ، (الفئة الثانية وتكرارها المناظر) ، .... وبيانيا فى شكل مساحات لمستطيلات (المدرج التكرارى) أو مساحة المضلع التكرارى أو مساحة المنحنى التكرارى .

جـ) أن الرسم البيانى للتوزيع التكرارى التجميعى المئوى هـو علاقة بين خط اعداد مفردات المجتمـع ككـل فـى شـكل متباينت والتكرار التجميعي المئوى المناظر لهذه المتباينات ، ومن ثم تظهر هذه العلاقة جدوليا فى شكل تباينـات لخـط اعداد مفردات المجتمع ككل والتكـرار التجميعـي المئـوى المناظر ، وبيانيا فى شكل مساحات ترسم بمعلومية اطـوال المتباينات والتكرار التجميعي المئوى المناظر ووفقا لمنـهج المنحنى التكرارى .

٢- من العرض الجدولى البيانى للتوزيع التكرارى التجميعى المئوى يتضح :-



أ) أن ٥ % من مجموع التكرارات (أي مجموع السائحين وهو ٢٠٠) انفاقهم اقل من ١٠ دولار ، وهذا ما يوضحة المنحنى المتجمع المئوى الصاعد . وبالتالى فان ٩٥ % من مجموع هذه التكرارات انفاقهم اكبر من ١٠ دولارات ، وهذا ما يوضحة المنحنى المتجمع المئوى الهابط .

ب) أن ٣٤ % من مجموع انفاقهم اقل من ٢٠ دولار ، وهذا ما يوضحة المنحنى المتجمع المئوى الصاعد . وبالتالى فان ٦٦ % من مجموع التكرارات انفاقهم اكبر من ٢٠ دولار ، وهذا ما يوضحة المنحنى المتجمع المئوى الهابط .

جـ) أن ٩٠,٥ % من مجموع انفاقهم اقل من ٣٠ دولار ، وهذا ما يوضحة المنحنى المتجمع المئوى الصاعد . وبالتالى فان ٩,٥ % من مجموع التكرارات انفاقهم اكبر من ٣٠ دولار ، وهذا ما يوضحة المنحنى المتجمع المئوى الهابط .

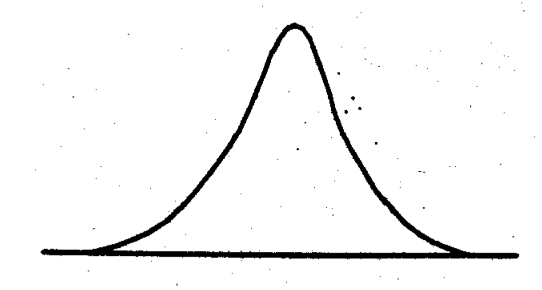
د) أن ٩٩ % من مجموع التكرارات (المجتمع) انفاقهم اقل من ٣٥ دولار ، وهذا ما يوضحة المنحنى المتجمع المئوى الصاعد . وبالتالى فان ١ % من هذا المجتمع انفاقهم اكبر من ٣٥دولار ، وهذا ما يوضحة المنحنى المتجمع المئوى الهابط .

هـ) أن المنحنى التجميعى المئوى الصاعد ، والمنحنى التجميعى المئوى الهابط يتقاطعان فى نقطة تعتبر غاية فى الأهمية الإحصائية حيث تحدد أحد المقاييس الهامة للنزعة المركزية (الوسيط) والتى سيرد شرحها فى الفصل القادم .

و) انه ينتج من تقاطع المنحنين معا إيجاد منحنى تكرارى اكثر قدرة على وصف توزيع المجتمع حيث :



- إذا كان ٥٠ % من مفردات المجتمع موزعة على الفئات الوسطى فقط ، وان الـ ٥٠ % الباقية موزعة على الفئات العليا والدنيا * ، فانه يمكن وصف هذا التوزيع بأنه توزيع متماثل وهو ما يسمى بالتوزيع الطبيعي حيث يمثل معظم الظواهر في الحياة ، ومنحنى هذا التوزيع هو المنحنى الطبيعي وهو يشبه الناقوس كما يلي : -



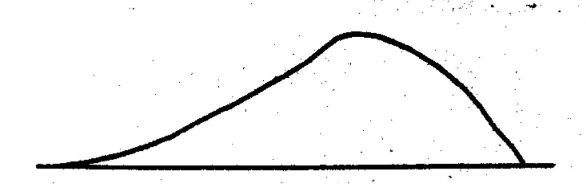
- إذا كان ٥٠ % فاكثر من مفردات المجتمع موزعة على الفئات العليا فقط ، وان الأقل من ٥٠ % الباقية موزعة على الفئات الوسطى والدنيا، فانه يمكن وصف هذا التوزيع بأنه توزيع ملتوى ناحية اليسار *حيث تقوم المفردات القليلة الباقية والموزعة على الفئات الوسطى والدنيا بسحب زيل المنحنى ناحية اليسار كما يلي : -

---

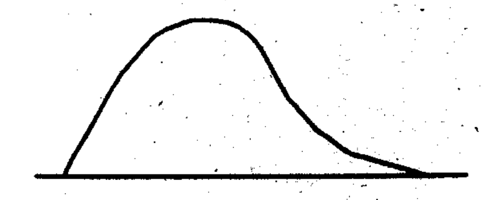
*التوزيع متناسب على الفئات

*اليسار هو يسار القارئ وليس يسار المنحنى





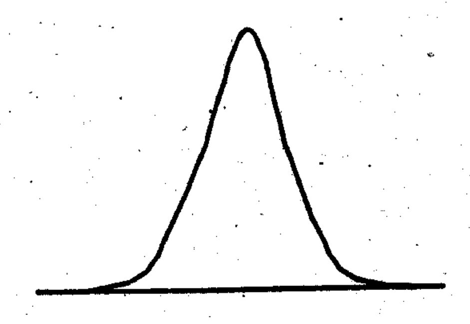
- إذا كان ٥٠ % فاكثر من مفردات المجتمع موزعة على الفئات الدنيا فقط ، وان الأقل من ٥٠ % الباقية موزعة على الفئات الوسطى والعليا ، فانه يمكن وصف هذا التوزيع بأنه توزيع ملتوى ناحية اليمين حيث تقوم المفردات القليلة الباقية والموزعة على الفئات الوسطى والعليا ا بسحب زيل المنحنى ناحية الجهة اليمين كما يلى : -



- إذا كان اكثر من ٥٠% من مفردات المجتمع موزعة على الفئات الوسطى فقط ، وان الأقل من ٥٠ % الباقية موزعة على الفئات الدنيا والعليا ، فانه يمكن وصف هذا التوزيع بأنه توزيع مدبب حيث الأكثر من ٥٠ % من مفردات الجتمع والموزعة على الفئات الوسطى فقط تجعل قمة



المنحنى ترتفع إلى اعلى من الوضع الطبيعى كما يلى : -



- إذا كان اقل من ٥٠ % من مفردات المجتمع موزعة على الفئات الوسطى فقط ، وان الأكثر من ٥٠ % الباقية موزعة على الفئات العليا والدنيا ، فانه يمكن وصف هذا التوزيع بأنه توزيع مفرطح حيث الأقل من ٥٠ % من مفردات المجتمع والموزعة على الفئات الوسطى فقط تجعل قمة المنحنى منخفضة لاسفل من الوضع الطبيعي كما يلى : -



ملحوظة : -

ولئن كان هذا الوصف المئوى لخصائص التوزيع يعد وصفا جيدا، إلا أن وصف خصائص التوزيع ما زال به شئ من العموم ، وان كنا نريد وصفا ذو تفاصيل اكثر ، وهذا ما سيتحقق فى الفصول القادمة حيث الوصف باستخدام القياس لخصائص التوزيع .

المجالات التطبيقية للتوزيع التكرارى التجميعى المئوى : -

يعتبر لورنز Lorenze الامريكى هو اول من استخدم فكرة التوزيع التكرارى التجميعى المئوى فى إيضاح ظاهرة تركيز توزيع الدخول على اصحاب الدخول ، أي درجة عدم المساواة فى توزيع الدخل بمعنى مدى انحراف التوزيع الفعلى للدخل عن التوزيع الأمثل ، والتوزيع الأمثل هو الذى يحقق التناسب التام بين عدد مكتسبي الدخل وبين مجموع الدخول التى يكتسبونها ، أي ان التوزيع الأمثل للدخل يتحقق إذا تحصل ٥٠ % من عدد مكتسبي الدخل على ٥٠% من مجموع الدخول أو إذا حصل ٢٠ % منهم على ٢٠ % من مجموع الدخول وهكذا ، ثم مقارنة التوزيع الفعلى بالتوزيع الأمثل وقياس الانحراف بينهما .

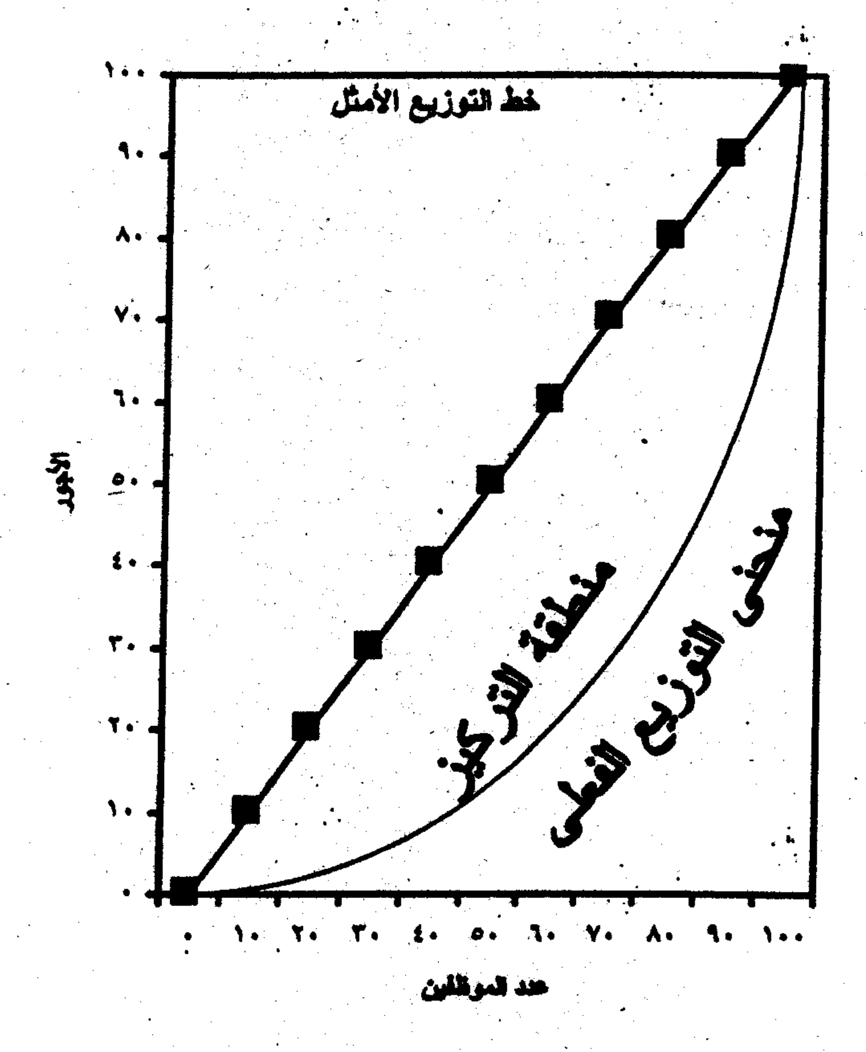


مثال

الجدول التالى هو التوزيع التكرارى التجميعي المئــوى الصــاعد لعــدد موظفى حكومة ما والأجور التى يتقاضونها فى سنة ما .

التوزيع الأمثل		التوزيع الفعلى	
تكرار تجميعى مئوى صاعد للعدد	تكرار تجميعى مئوى صاعد للأجور	تكرار تجميعى مئوى صاعد للعدد	تكرار تجميعى مئوى صاعد للأجور
صفر	صفر	صفر	صفر
١٠	١٠	١٠	١
٢٠	٢٠	٢٠	١,٣
٣٠	٣٠	٤٠	٣
		٥٠	٦
٥٠	٥٠	٧٠	١١
		٩٠	٣٥
٨٠	٨٠	٩٥	٦٠
		٩٨	٧٧
١٠٠	١٠٠	٩٩	٨٩
		١٠٠	١٠٠

المطلوب : قياس نسبة التركيز فى التوزيع الفعلى



بعد ما يتم الرسم البيانى للتوزيع الأمثل والفعلى كما فى الرسم فان نسبة التركيز Concentration Ratio هى المساحة المحصورة بين المنحنى وخط التوزيع الأمثل ، وكلما صغرت هذه المساحة كلما اقترب التوزيع الفعلى من التوزيع الأمثل ، وبقسمة المساحة المحصورة بين



المنحنى وخط التوزيع الأمثل على مساحة المثلث الواقعة فيه فانه ينتج لدينا ما يسمى بنسبة التركيز .

.. نسبة التركيز = المساحة الواقعة بين المنحنى وخط التوزيع الأمثل / مساحة المثلث الواقعة فيه

وهذه النسبة تنحصر بين الصفر والواحد الصحيح ، وكلما اقتربت من الصفر دل ذلك على قرب التوزيع الفعلي من التوزيع الأمثل والعكس صحيح .

الفصل الثاني

النزعة المركزية للبيانات الإحصائية

تمهيد : -

انتهينا مما سبق إلى جمع البيانات ثم تلخيصها فى شكل جداول ورسوم بيانية ، إلا أن هذا التلخيص غير كاف حيث لا يعطى وصفا كافيا عن مجموعة بيانات المتغير الإحصائي محل الدراسة ، لهذا ظهرت الحاجة إلى تلخيص مجموعة البيانات فى شكل افضل أي فى صورة رقم ومن ثم يمكن فهم البيانات بسهولة ويسر وأيضا إمكانية المقارنة بين مجموعتين أو اكثر من البيانات .

واستنادا إلى النزعة المركزية Central tendency للبيانات أي ميولها إلى التركز أو التراكم حول نقطة حيث تعرف بالنقطة المتوسطة ، فقد تم استغلال هذه النقطة المتوسطة فى معرفة القيمة المتوسطة Average عندها ومن ثم يمكن معرفة خصائص هامة عن مجموعة بيانات المتغير محل الدراسة .

ويتم قياس القيمة المتوسطة أي النزعة المركزية بعدة مقاييس تعرف بمقاييس المتوسط أو بمقاييس النزعة المركزية ، ومن أهم تلك المقاييس : -

١. الوسط الحسابى Arithmetic mean

٢. الوسط الهندسى Geometric mean

٣. الوسط التوافقى Harmaic mean

٤. الوسيط Median

٥. المنوال Mod

ويراعى فى مجموعة بيانات المتغير الإحصائي التى يراد قياس القيمة المتوسطة لها ما يلى : -

١. أن تكون مجموعة البيانات متجانسة (والمقصود بالتجانس هو التجانس المقبول وليس التجانس التام) ، فلا يجوز مثلا حساب متوسط الدخل لعدد من العمال بما فيهم المدير ، حيث أن دخل المدير يختلف كثيرا عن دخل العمال الذين يرأسهم . بينما يجوز حساب المتوسط لسن العمال والمدير معا ، حيث التقارب فى السن (التجانس المقبول) .

٢. إذا كانت مجموعة بيانات المتغير الإحصائي كبيرة فلابد أن يكون تبويبها (وضعها فى جدول توزيع تكرارى) سليما ، حيث التبويب غير السليم يؤدى إلى حساب قيمه متوسطة مضللة .

كما يراعى فى المقياس المستخدم ما يلى : -

١. أن يعتمد فى حسابه على جميع مفردات المتغير الإحصائي محل الدراسة .

٢. ألا تتأثر قيمتة باختيار العينة المسحوبة من المجتمع .

٣. أن يكون سهلا فى حسابه بعيدا عن الافكار الرياضية المعقدة ، على ألا يعنى ذلك التضحية بالدقة فى سبيل سهولة الحساب .



٤. أن تكون قيمته محددة ولا تحتمل التأويل ومحسوبة طبقا لمعادلة رياضية ثابتة وذلك حتى تكون قيمته موضوعية لا يختلف اثنان في تقديرها .

كما يراعى عند اختيار أيا من مقاييس المتوسط ما يلي : -

١. توزيع البيانات .

٢. وجود قيم شاذه من عدمه

٣. الغرض من استخدام المقياس هل هو للقياس فقط أم سيستخدم في تقدير حسابات أخرى .

٤. نوع القياس لبيانات المتغير الإحصائي هل هي قياسات مباشرة أم تم لها تحويلات رياضية معينة كأن تكون على صورة نسب أو جذور أو لوغاريتمات ..الخ

مقاييس النزعة المركزية في حالة المجتمع الإحصائي الصغير : -

إذا كان المجتمع الإحصائي صغير أي هو متغير لدرجات نجاح ٧ طلاب في مادة الإحصاء {٦٠ ، ٤٠ ، ٣٠ ، ٧٠ ، ٥٠ ، ٩٠ ، ١٠} ، فاحسب القيمة المتوسطة لمجموعة بيانات هذا المتغير .

الحل

١- الوسط الحسابي : -

يعرف الوسط الحسابي بأنه أحد مقاييس النزعة المركزية وأكثرها استخداما ، وهو القيمة التي تمثل قيم مجموعة بيانات المتغير الإحصائي للتمركز حولها ، ويحسب عن طريق حاصل جمع قيم مجموعة المفردات مقسوما على عددها .



$$\bar{س} = \frac{مجـ س}{ن}$$

حيث :

$\bar{س}$: هى رمز للوسط الحسابى (المتوسط الحسابى)*

مجـ س : هى مجموع قيم المفردات .

ن : هى عدد المفردات

$$\therefore \bar{س} = \frac{٦٠ + ٤٠ + ٣٠ + ٧٠ + ٥٠ + ٩٠ + ١٠}{٧}$$

$$= \frac{٣٥٠}{٧}$$

$$= ٥٠$$

٢- الوسط الهندسي :-

يعرف الوسط الهندسي بأنه الجذر النوني لحاصل ضرب قيم مجموعة مفردات المتغير الإحصائي محل الدراسة .

فإذا كانت قيم مفردات متغير إحصائي ما هى $س_١$ ، $س_٢$ ،.... $س_ن$

فإن الوسط الهندسي $س_هـ$ =

$$= \sqrt[ن]{س_١ \times س_٢ \times ....\times س_ن}$$

*نظرا لكثرة استعمال الوسط الحسابى فى قياس القيمة المتوسطة فقد أطلق عليه المتوسط الحسابى .



وبتطبيق ذلك على المثال السابق يكون :

الحـــل

$$\therefore \text{هـ} = \sqrt[\text{ن}]{\text{س}_{١} \times \text{س}_{٢} \times .... \times \text{س}_{\text{ن}}}$$

$$\therefore \text{هـ} = \sqrt[٧]{٦٠ \times ٤٠ \times ٣٠ \times ٧٠ \times ٥٠ \times ٩٠ \times ١٠}$$

$$= \sqrt[٧]{٢٢٦٨٠٠٠٠٠٠٠٠}$$

$$= ٤١,٩$$

ونظرا لصعوبة إجراء العمليات الحسابية هذه خاصة إذا ما كبرت قيم المفردات فقد استخدمت اللوغاريتمات لتسهيل العمليـــات الحســـابية ، وبتطبيق ذلك على المثال السابق يكون : -

الحـــل

$$\text{لو هـ} = \frac{١}{\text{ن}} (\text{لو س}_{١} + \text{لوس}_{٢} + ... + \text{لو س}_{\text{ن}})$$

$$= \frac{\text{مجـ لوس}}{\text{ن}}$$

$$= \frac{\text{لو}٦٠ + \text{لو}٤٠ + \text{لو}٣٠ + \text{لو}٧٠ + \text{لو}٥٠ + \text{لو}٩٠ + \text{لو}١٠}{٧}$$

$$= \frac{١,٧٨ + ١,٦ + ١,٤٨ + ١,٨٥ + ١,٧ + ١,٩٥ + ١}{٧}$$

$$= \frac{١١,٣٦}{٧}$$

$$= ١,٦٢$$

.. هـ = ٤١,٩ وذلك بإيجاد العدد المقابل للقيمة لو ١,٦٢ .

ويتضح أنها نفس النتيجة السابقة

٣- الوسط التوافقى :-

ويعرف بأنه مقلوب الوسط الحسابى لمقلوب القيم .

فإذا كانت قيم مفردات متغير إحصائي ما هى $س_١$، $س_٢$ ، ...... $س_ن$

$$\text{فإن الوسط التوافقى (ق)} = \frac{ن}{مج \frac{١}{س}}$$

وبتطبيق ذلك على المثال السابق يكون : -

الحـــل

$$.. ق = \frac{ن}{مج \frac{١}{س}}$$

$$= \frac{٧}{\frac{١}{٦٠} + \frac{١}{٤٠} + \frac{١}{٣٠} + \frac{١}{٧٠} + \frac{١}{٥٠} + \frac{١}{٩٠} + \frac{١}{١٠}}$$

$$= \frac{٧}{٠,٢٢}$$

= ٣١,٨٢

٤- الوسيط : -

ويعرف بأنه القيمة التى تتوسط مجموعة المفردات بعد ترتيبها ترتيبا تصاعديا أو تنازليا ، ولايجاد قيمة الوسيط (ط) يتعين اتباع الاتى: -

١. ترتيب القيم تصاعديا أو تنازليا .

٢. تحديد موقع الوسيط .

٣. استخراج قيمة الوسيط

وبتطبيق ذلك على المثال السابق يكون : -

الحل

.. قيم مجموعة مفردات المتغير هى :

٦٠ ، ٤٠ ، ٣٠ ، ٧٠ ، ٥٠ ، ٩٠ ، ١٠

.. بترتيبها تصاعديا تكون :

١٠ ، ٣٠ ، ٤٠ ، ٥٠ ، ٦٠ ، ٧٠ ، ٩٠

، ولتحديد موقع الوسيط يلاحظ أن عدد المفردات ٧ أي فرديا لذلك يتحدد موقع الوسيط بالقانون $\frac{ن + ١}{٢}$



.. موقع الوسيط = $\frac{٧ + ١}{٢}$

= $\frac{٨}{٢}$

= ٤

أي أن الموقع هو المفردة ذات الترتيب الرابع

، ولاستخراج قيمة الوسيط يتم ملاحظة قيمة المفردة ذات الترتيب الرابع تكون هى قيمة الوسيط

.. ط = ٥٠

ملحوظة : -

ماذا لو كان عدد مفردات المتغير الإحصائي عدد زوجى

مثال

إذا كانت قيم مفردات مجموعة متغير إحصائي ما هى :

{ ٤٦٥ ، ٣٧٥ ، ٦١٢ ، ٤٦٩ ، ٥١٦ ، ٧٢١ ، ٤٠٢ ، ٣٢٦ }

فاحسب الوسيط .

الحــــل

.. مجموعة قيم المفردات هى :



{ ٤٦٥ ، ٣٧٥ ، ٦١٢ ، ٤٦٩ ، ٥١٦ ، ٧٢١ ، ٤٠٢ ، ٣٢٦ } .

.. ترتيبها تنازليا تكون :

{ ٧٢١ ، ٦١٢ ، ٥١٦ ، ٤٦٩ ، ٤٦٥ ، ٤٠٢ ، ٣٧٥ ، ٣٢٦ }

، ولتحديد قيمة الوسيط يلاحظ أن عدد المفردات ٨ أي عدد زوجي لذلك يتحدد موقع الوسيط من خلال موقعين يتحددان بالقانونين :

$\frac{ن}{٢}$ ، $\frac{ن}{٢}$ + ١ وعليه فموقع الوسيط يتحدد بتحديد موقع أول وموقع ثاني .

.. الموقع الأول أى $\frac{ن}{٢}$ أى $\frac{٨}{٢}$ = ٤ ، الموقع الثانى أى $\frac{ن}{٢}$ +١ أى $\frac{٨}{٢}$ + ١ = ٥

.. موقع الوسيط يقع بين المفردة ذات الترتيب الرابع والمفردة ذات الترتيب الخامس .

، ولاستخراج قيمة الوسيط يتم ملاحظة قيمة المفردة ذات الترتيب الرابع وقيمة المفردة ذات الترتيب الخامس فتكون قيمة الوسيط هى القيمة التى تتوسطهما .

$$.. ط = \frac{٤٦٩ + ٤٦٥}{٢}$$

$$= ٤٦٧$$

٥- المنوال : -

ويعرف بأنه القيمة الأكثر شيوعا أو تكرارا فى مجموعة بيانـــات المتغير الإحصائي موضوع الدراسة ، وبتطبيق ذلك على المثال الســـابق والمتناول فى المقايس السابقة يكون : -

الحـــل

.. قيم مجموعة المفردات هى :

{ ٦٠ ، ٤٠ ، ٣٠ ، ٧٠ ، ٥٠ ، ٩٠ ، ١٠ } .

.. بملاحظة قيم مجموعة المفردات فلا توجد مفردة قد تكررت .

.. مجموعة المفردات هذه لا يوجد بها قيمة منوالية .

مثال آخـــر

إذا كانت مجموعة قيم مفردات متغير إحصائي ما هى :

{ ٦٠ ، ٤٠ ، ٣٠ ، ٧٠ ، ٥٠ ، ٩٠ ، ١٠ }

فاحسب المنوال (م)

الحـــل

.. قيم مجموعة المفردات هى :

{ ٦٠ ، ٤٠ ، ٣٠ ، ٧٠ ، ٦٠ ، ٥٠ ، ٩٠ ، ١٠ } .



∴ م = ٦٠ حيث قيمة المفردة الأكثر تكرارا

مثـــال تمهيـــدى

إذا كانت مجموعة قيم مفردات متغير إحصائي ما هي :

{ ٤ ، ١٤ ، ١٥ ، ٢٩ ، ٣ ، ١٤ ، ٣٤ ، ٩ ، ١٤ ، ١٩ ، ٣٥ ، ١٧
، ١٣ ، ٢٤ ، ٢٤ ، ٩ ، ١٤ ، ٢٣ ، ٢٨ ، ١٨ }

فاحسب مقاييس النزعة المركزية لهذه البيانات .

الحـــل

$$\bar{س} = \frac{مج\, س}{ن}$$

$$= \frac{٤+١٤+١٥+٢٩+٣+١٤+٣٤+٩+١٤+١٩+٣٥+١٧+١٣}{٢٠}$$

$$\frac{٢٤+٢٤+٩+١٤+٢٣+٢٨+١٨}{}$$

= ١٨

ونظرا لصعوبة العمليات الحسابية والتى ستزداد صعوبة إذا ما كان عدد المفردات كبير أي ٥٠ مفردة مثلا أو ١٠٠ أو ١٠٠٠ ، فهنا يستحيل اتباع نفس الخطوات السابقة فى قياس القيمة المتوسطة باستخدام مقاييس النزعة المركزية ، وعلية كان لابد من التوصل إلى أسلوب آخر يسهل عملية الحساب .

انه لحساب القيمة المتوسطة لمجموعة بيانات مجتمع إحصائي كبير كما فى المثال { ١، ١٠، ... ، ١٧ } فلابد من تفريغ هذه البيانات فى جدول توزيع تكرارى كما سبق ، ومن ثم نجد التوزيع التكرارى هو .

جدول توزيع تكرارى لمجموعة بيانات مجتمع إحصائي

عبارة عن درجات امتحان ٢٠ طالب فى مادة الإحصاء

الفئات	(٠، ١٠)	(١٠ – ٢٠)	(٢٠ – ٣٠)	(٣٠ – ٤٠)	المجموع
التكرارات	٤	٩	٥	٢	٢٠

والقياس القيمة المتوسطة لهذه البيانات نستخدم مقاييس النزعة المركزية كما يلى : -

١- الوسط الحسابى : -

تكوين الجدول الإحصائي اللازم : -

الفئات	التكرار(ك)	مركز الفئة (س)	س × ك
(٠ – ١٠)	٤	٥	٢٠
(١٠ – ٢٠)	٩	١٥	١٣٥
(٢٠ – ٣٠)	٥	٢٥	١٢٥
(٣٠ – ٤٠)	٢	٣٥	٧٠
المجموع	٢٠		٣٥٠

ثم نستخدم القانون التالى : -

$$\bar{س} = \frac{مجـ\ س\ ك}{مجـ ك}$$

حيث :

$\bar{س}$: ترمز له سط الحسابى .



مجـ س ك : مجموع حاصل ضرب عناصر العمود ك في نظائرها من عناصر العمود س

مجـ ك : مجموع التكرارات

$$\therefore \bar{س} = \frac{٣٥٠}{٢٠}$$

$$= ١٧,٥$$

٢- الوسط الهندسي : -

تكوين الجدول الإحصائي اللازم :-

الفئات	التكرار(ك)	مركز الفئة (س)	لوس	ك × لوس
(٠ – ١٠)	٤	٥	٠,٧٠	٢,٨
(١٠ – ٢٠)	٩	١٥	١,١٨	١٠,٦٢
(٢٠ – ٣٠)	٥	٢٥	١,٤٠	٧,٠
(٣٠ – ٤٠)	٢	٣٥	١,٥٤	٣,٠٨
المجموع	٢٠	/	/	٢٣,٥

ثم نستخدم القانون التالي : -

$$لو\ هـ = \frac{مجـ\ ك\ لوس}{مجـ ك}$$

حيث :

لو هـ : لوغاريتم الوسط الهندسي هـ



مجـ ك لو س : مجموع حاصل ضرب عناصر العمود ك في نظائرها من عناصر العمود لو س

مجـ ك : مجموع التكرارات

$$\therefore \text{لو هـ} = \frac{٢٣,٥}{٢٠}$$

$= ١,١٧٥$

∴ هـ = ١٤,٩٦ وذلك بإيجاد العدد المقابل لقيمة لو هـ

٣- الوسط التوفقى : -

تكوين الجدول الإحصائي اللازم : -

الفئات	التكرار = (ك)	مركز الفئة (س)	$\frac{١}{س}$	ك × $\frac{١}{س}$
(٠ – ١٠)	٤	٥	$\frac{١}{٥}$	٠,٨٠
(١٠ – ٢٠)	٩	١٥	$\frac{١}{١٥}$	٠,٦٠
(٢٠ – ٣٠)	٥	٢٥	$\frac{١}{٢٥}$	٠,٢٠
(٣٠ – ٤٠)	٢	٣٥	$\frac{١}{٣٥}$	٠,٠٦
المجموع	٢٠	/	/	١,٦٦

ثم نستخدم القانون التالى : -

$$\text{ق} = \frac{\text{مجـ ك}}{\text{مجـ ك} \cdot \frac{١}{س}}$$



حيث :

ق : هو رمز الوسط التوافقى

مجـ ك : مجموع عدد التكرارات

مجـ ك . $\frac{١}{س}$: مجموع عدد حاصل ضرب عناصر العمــود ك فى نظائرها من عناصر $\frac{١}{س}$ ، $\frac{١}{س}$ هى مقلوبات عمود مركز الفئات (س) .

$$\therefore ق = \frac{٢٠}{١,٦٦}$$

= ١٢,٠٥ درجة

٤- الوسيـــط : -

تكوين الجدول الإحصائي اللازم : -

الفئات	التكرار(ك)	مركز الفئة (س)	لوس
		الحدود العليا للفئات	تكرار متجمع صاعد
(٠ - ١٠)	٤	١٠-	٤
(١٠ - ٢٠)	٩	٢٠-	١٣
(٢٠ - ٣٠)	٥	٣٠-	١٨
(٣٠ - ٤٠)	٢	٤٠-	٢٠
المجموع	٢٠	/	/

ثم نتبع الخطوات التالية لاستخراج قيمة الوسيط :

أ - نحدد الفئة الوسيطية من الجدول التكرارى المتجمع الصاعد كما يلى :



- نحدد تكرار الوسيط بالقانون $\frac{مجـ ك}{٢}$ * وذلك من على عمود التكرار المتجمع الصاعد .

$$\therefore \text{تكرار الوسيط} = \frac{٢٠}{٢}$$

$$= ١٠$$

- نحدد الفئة الوسيطية وهى الفئة التى تقابل تكرار الوسيط

∴ الفئة الوسيطيه هى (١٠ – ٢٠)

ب- نستخرج قيمة الوسيط من القانون التالى :-

$$ط = ب + \frac{ت_١ \times ل}{ت_١ + ت_٢}$$

حيث :

ط : هى رمز القيمة الوسيطية

ب : هى بداية الفئة الوسيطية

$ت_١$: هى الفرق * بين تكرار الوسيط والتكرار السابق عليه

$ت_٢$: هى الفرق بين تكرار الوسيط والتكرار التالى عليه

ل : هى طول الفئة الوسيطية

$$\therefore ط = ١٠ + \frac{(١٠ - ٤) \times ١٠}{(١٠ - ٤) + (١٠ - ١٣)}$$

---

* على أساس أن ترتيب الوسيط هو نصف مجموع التكرارات .

* المقصود بالفرق هو الفرق المطلق أي بصرف النظر عن الإشارة

$$= ١٠ + \frac{٦٠}{٦ + ٣}$$

$$= ١٦,٧$$

٥- المنوال : -

تكوين الجدول الإحصائي اللازم : -

الفئات	التكرارات (ك)
(٠ – ١٠)	٤
(١٠ – ٢٠)	٩
(٢٠ – ٣٠)	٥
(٣٠ – ٤٠)	٢
المجموع	٢٠

ثم نتبع الخطوات التالية لاستخراج قيمة المنوال :

أ – نحدد الفئة المنوالية من الجدول التكرارى العادى كما يلى :

– نحدد تكرار المنوال من القانون بأنه الأكبر تكرار وذلك على عمود التكرار العادى .

.. تكرار المنول = ٩

– نحدد الفئة المنوالية وهى الفئة التى تقابل اكبر تكرار

.. الفئة المنوالية هى (١٠ – ٢٠)

ب – نستخرج قيمة المنول من القانون التالى :

$$\text{م} = \text{ب} + \frac{\text{ت}_{١} \times \text{ل}}{\text{ت}_{١} + \text{ت}_{٢}}$$

حيث :

م : هى رمز القيمة المنوالية

ب : هى بداية الفئة المنوالية

ت ١ : هى الفرق بين تكرار الفئة المنوالية والتكرار السابق عليه

ت ٢ : هى الفرق بين تكرار الفئة المنوالية والتكرار التالى عليه

ل : هى طول الفئة المنوالية

$$\therefore \text{ط} = ١٠ + \frac{(٩ - ٤) \times ١٠}{(٩ - ٤) + (٩ - ٥)}$$

$$= ١٠ + \frac{٥ \times ١٠}{٥ + ٣٤}$$

$$= ١٥,٦$$

خواص مقاييس النزعة المركزية : -

بالنسبة للوسط الحسابى : -

١. مجموع انحرافات القيم من متوسطها الحسابى يساوى صفر .

الإثبـــــــات

سبق حساب المتوسط الحسابى ($\bar{س}$) لمجموع البيانات { ٦٠، ٤٠ ، ٣٠ ، ٧٠ ، ٥٠ ، ٩٠ ، ١٠ } وتبين انه يساوى ٥٠ .

ولحساب مجموع انحرافات قيم مفردات المجموعة عن متوسطها الحسابى $\bar{س}$ والذى يساوى ٥٠ فان : -

مج (س-$\bar{س}$) = [(٦٠ - ٥٠) + (٤٠ - ٥٠) + (٣٠ - ٥٠) + (٥٠ - ٥٠) + (٩٠ - ٥٠) + (١٠ - ٥٠)]

= صفر

هذا ولو تم حساب انحرافات هذه المفردات عن أي قيمة أخـــرى خـــلاف المتوسط $\bar{س}$ (٥٠) ، فان مجموع هذه الانحرافات يختلف عن الصفر .

كما يمكن اثبات ذلك جبريا ، فإذا رمزنا إلى الانحرافات السابقة بالرمز $ح_١$ ، $ح_٢$ ، $ح_٣$ ، $ح_٤$ ، $ح_٥$ ، $ح_٦$ ، فان :

$ح_١$ = $س_١$ - $\bar{س}$

$ح_٢$ = $س_٢$ - $\bar{س}$

$ح_٣$ = $س_٣$ - $\bar{س}$

" " "

" " "

$ح_٦$ = $س_٦$ - $\bar{س}$

.. مج ح = مج س - ن $\bar{س}$

$$= \text{مج س} - \text{ن} \times \frac{\text{مج س}}{\text{ن}}$$

$$= \text{مج س} - \text{مج س}$$

= صفــــــــر

٢. مجموع مربعات انحرافات القيم عن وسطها الحسابى اقل مـــا يمكن ، أي اقل من مجموع مربعات الانحرافات عن أي قيمـــة أخرى ، وسيرد شرح هذه الخاصية فى الفصل القادم لتعلقـــها بموضوع التشتت .

٣. إذا أضفنا أو طرحنا مقدار ثابت لكل قيم المفردات ، فان الوسط الحسابى الجديد يساوى الوسط الحسابى للقيم الأصلية للمفردات مضافا إلية أو مطروحا منه نفس المقدار الثابت .

٤. إذا ضربنا أو قسمنا كل قيم المفردات على مقدار ثابت ، فـــان الوسط الحسابى الجديد يساوى الوسط الحسابى للقيم الأصليـــة للمفردات مضروبة أو مقسومة على نفس المقدار الثابت .

٥. يتأثر الوسط الحسابى بالقيم المتطرفة فى المجموعـــة لذلـــك لا يفضل استخدامه إذا كان بمجموعة البيانات قيم متطرفة (شاذة)

٦. يتأثر بكل مفردات المجموعة لذلك يتم استخدامه فى التوزيـــع القريب من التماثل ، أما إذا استخدم فى التوزيع غير المتمـــاثل بدرجة كبيرة فانه يعطى قيمة مضللة .

٧. لا يمكن استخدامه فى حالة الجداول التكرارية المفتوحة نظـــرا لعدم إمكان إيجاد مركز الفئة المفتوحة ومن ثم إهدار هذه الفئـــة فى عملية حسابه .



٨. إذا تم تقسيم مجموعة المفردات إلى مجموعات جزئية ، فانه خطأ أن يتم حساب المتوسط الحسابي لمجموعة المفردات على انه يساوى متوسط متوسطات المجموعات الجزئية دون الترجيح بأوزان هذه المجموعات

مثال

نفرض انه تم تقسيم مجموعة درجات السبعة طلاب فى المثال السابق إلى المجموعين { ٦٠ ، ٤٠ ، ٣٠ } ، { ٧٠ ، ٥٠ ، ٩٠ ، ١٠ } .

فانه خطأ أن يتم الحساب كما يلى : -

$$\bar{س}\ \text{للمجموعة الأولى} = \frac{٦٠ + ٤٠ + ٣٠}{٣} = ٤٣,٣$$

$$\bar{س}\ \text{للمجموعة الثانية} = \frac{٧٠ + ٥٠ + ٩٠ + ١٠}{٤} = ٥٥$$

$$\therefore \bar{\bar{س}}\ \text{العام} = \frac{٤٣,٣ + ٥٥}{٢} = ٤٩,١٥$$

و إنما لابد من الترجيح بأوزان المجموعات كما يلى : -

$$\bar{\bar{س}}\ ع = \frac{٤٣,٣ \times ٣ + ٥٥\times٤}{٣ + ٤} = ٥٠$$

بالنسبة للوسط الهندسي : -

١. الوسط الهندسي لمجموعة المفردات دائما اقل من وسطها الحسابي .



٢. لا يمكن حساب الوسط الهندسي إذا كانت مجموعة المفـــردات إحداها ذات قيمة صفر أو بالسالب . ذلك انه أن كانت إحداهـــا صفر فان الوسط الهندسي سيساوى الصفر ، وان كانت إحداهـا بالسالب فان الوسط الهندسي سيكون كمية تخيلية (غير حقيقية)

٣. يستخدم فى الظواهر التى تتبع فى تغيرها قانون الربح المركب ومن ثم يكون خطأ استخدام الوسط الحسابى فى هذه الحالة .

مثـــال

إذا كان عدد سكان فى سنة ١٩٢٧ يساوى ٩ مليون نسمة ، وفـــى عام ١٩٣٧ يساوى ١٦ مليون نسمة فاحسب متوسط عدد السكان فـى تلك الفترة .

الحـــل

يكون خطئنا إذا تم الحساب كما يلى : -

الوسط الحسابى لعدد السكان خلال الفترة = $\frac{٩ + ١٦}{٢}$ = ١٢,٥ مليـــون نسمة والسبب أن ظاهرة السكان تتبع فى تغيرها قانون الربـــح المركـــب وليس الربح البسيط ولذلك يستخدم الوسط الهندسى .

الوسط الهندسي كقيمة متوسطة لعدد السكان خلال الفترة =

= $\sqrt{٩ \times ١٦}$ = ١٢ مليون نسمة .

الإثبـــات

∴جـ = أ (١ + ر)^{ن}

حيث جـ المبلغ النهائى ، أ هى المبلغ الأصلي ، ر سعر الفائدة السنوى ، ن عدد السنوات .

وبتطبيق هذا القانون على ظاهرة السكان في المثال .

.. ١٦ = ٩ (١ + ر)^{١٠}

.. ر = ٠,٠٥٩١

.. معدل الزيادة السنوى فى عدد سكان = ٥,٩١ %

.. عدد السكان فى متوسط الفـــترة = ٩ (١ + ٠,٠٥٩١)^{٥}

= ١٢ مليون نسمة

٤. يستخدم الوسط الهندسي كثيرا فى إيجاد متوسط التغير النسبي للاسعار عند عمل الأرقام القياسية ويفضل عـن المتوسطات الأخرى .

بالنسبة للوسط التوافقي : -

لا يستعمل إلا نادرا حيث لا تتوفر فيه معظم الشروط الواجب توافرها في المقياس الجيد .

بالنسبة للوسيط: -

١. الوسيط مقياس لمتوسط رتب القيم وليس لمتوسط القيم .

٢. لا يتأثر بالقيم الشاذة أو المتطرفة ، لذلك يفضل الوسيط فى قياس النزعة المركزية لمجموعة البيانات إذا كان بها قيمة أو قيم قليلة متطرفة أو التوزيع ملتوى التواء شديد .

٣. لا يتأثر بكون الجدول التكرارى مفتوح حيث لا يتطلب حسابه إيجاد مراكز الفئات .

المنـــوال :-

١. لا يتأثر بالقيم الشاذة ولا بالجداول المفتوحـــة ومـــن ثـــم فـــهو كالوسيط فى هذه الخاصية .

٢. لا يمكن الاعتماد عليه إذا كان عدد المفردات كبير .

استخدام مقاييس النزعة المركزية فى وصف خصائص التوزيع :-

١. إذا كان $\bar{س} = ط = م$ فان التوزيع التكرارى متماثل

مثال

الجدول التالى هو التوزيع التكرارى لامتحان عدد ١٢٨ طالب في مادة الاقتصاد .

الفئات	(٠-١٠)	(١٠-٢٠)	(٢٠-٣٠)	(٣٠-٤٠)	(٤٠-٥٠)	(٥٠-٦٠)	(٦٠-٧٠)	(٧٠-٨٠)	المجموع
التكرارات	١	٧	٢١	٣٥	٣٥	٢١	٧	١	١٢٨

والمطلوب :-

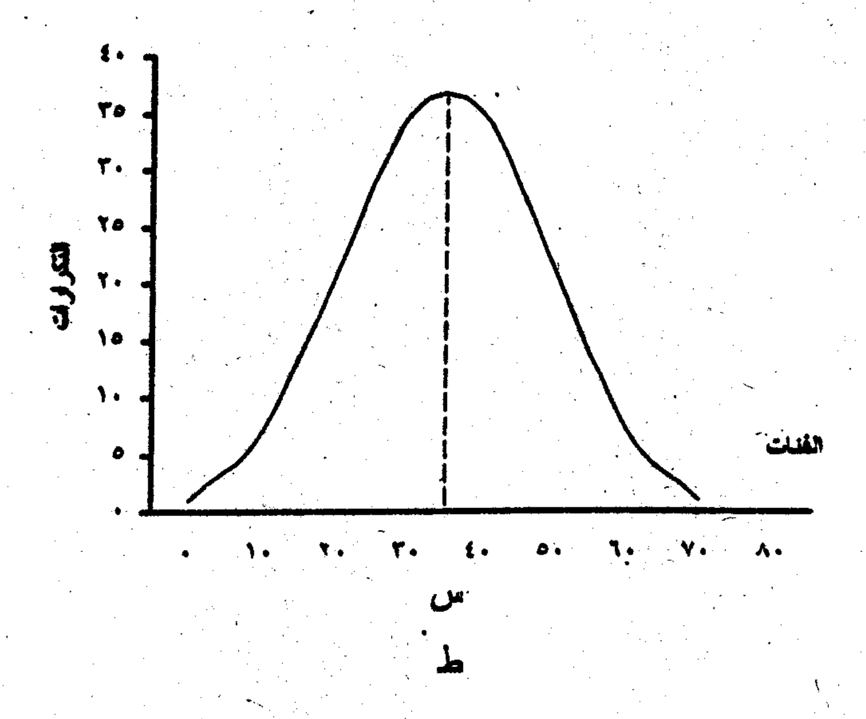
حساب كل من $\bar{س}$ ، ط ، م ثم إيجاد العلاقة بينهم واســـتخدام هـــذه العلاقة فى وصف التوزيع .

الحـــل

بحساب كل من $\bar{س}$ ، ط ، م تبين أن :-

$\bar{س} = ٤٠$ ، ط = ٤٠ ، م = ٤٠

وعليه فالتوزيع متماثل ويتضح ذلك بيانيا فى الشكل التالى : -



٢- إذا كان س < ط < م فان التوزيع ملتوى يسارا .

مثال

الجدول التالى هو التوزيع التكرارى لامتحان عدد ١٢٨ طالب فى مـــادة الرياضيات

الفئات	(٠-١٠)	(١٠-٢٠)	(٢٠-٣٠)	(٣٠-٤٠)	(٤٠-٥٠)	(٥٠-٦٠)	(٦٠-٧٠)	(٧٠-٨٠)	المجموع
التكرارات	٢	٦	١٠	١٥	٢٠	٤٠	٢٥	١٠	١٢٨

والمطلوب : -

حساب كل من س ، ط ، م ثم إيجاد العلاقة بينهم واستخدام هـــذه العلاقة فى وصف التوزيع .

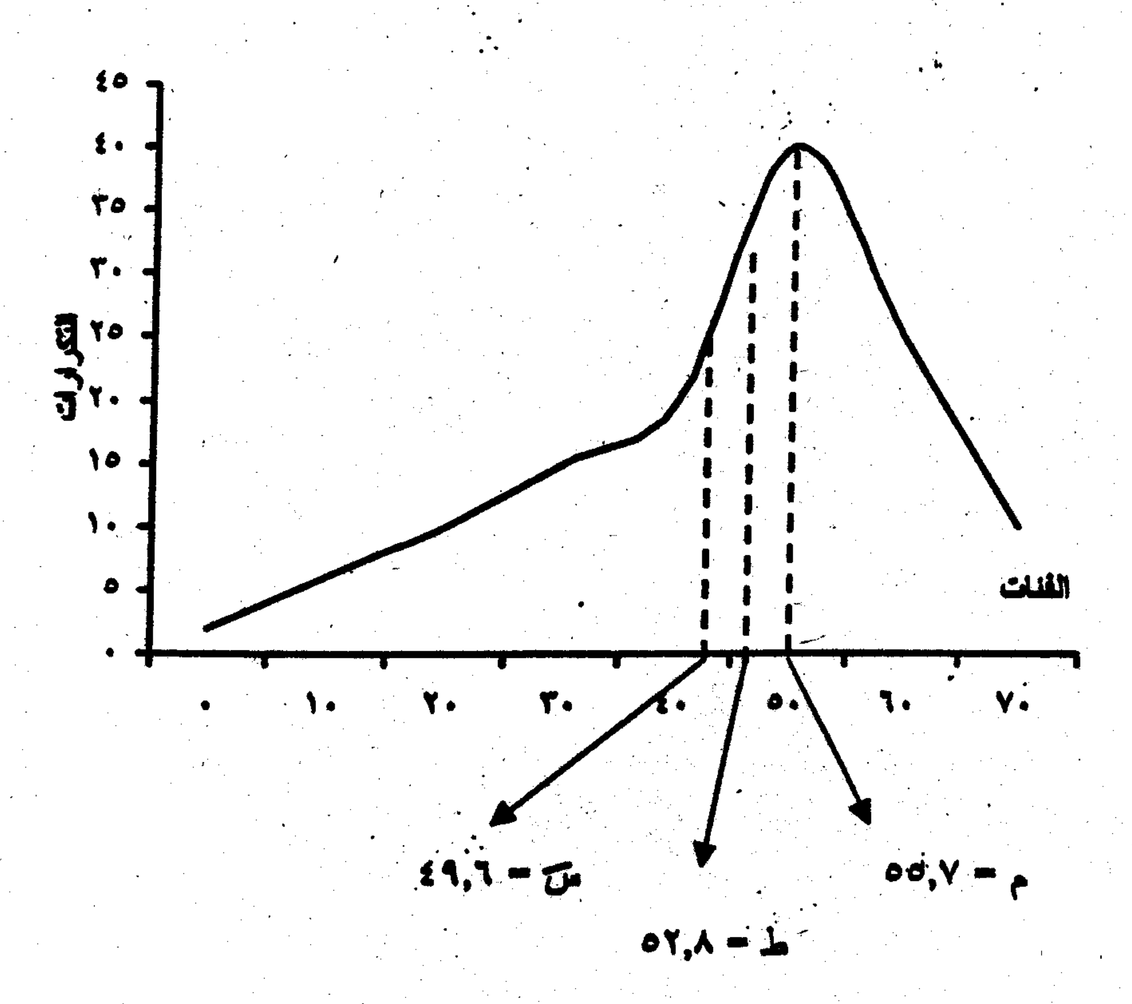


الحـــل

بحساب كل من سَ ، ط ، م تبين أن : -

سَ = ٤٩,٦ ، ط = ٥٢,٨ ، م = ٥٥,٧

وعليه فالتوزيع ملتوى يسارا ، ويتضح ذلك بيانيا فى الشكل التالى : -



٣- إذا كان سَ < ط < م فان التوزيع ملتوى يمينا.

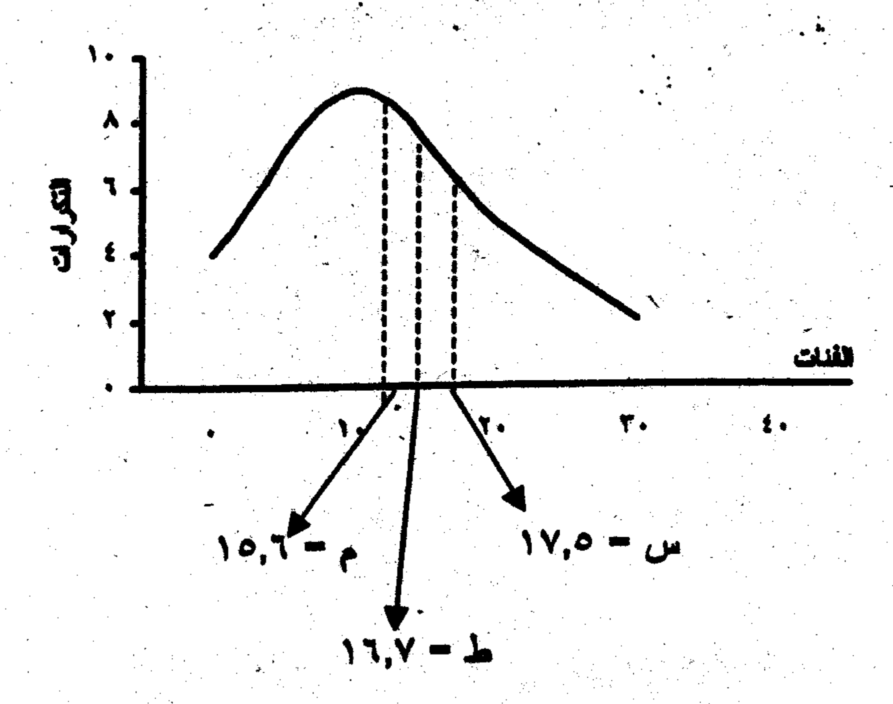
مثـــــال

عند قياس النزعة المركزية في المثال عن مجتمع امتحان ٢٠ طالب فـــى مادة الإحصاء تبين أن :

س̄ = ١٧,٥ ، ط = ١٦,٧ ، م = ١٥,٦

∴ س̄ > ط > م

وعليه فالتوزيع ملتوى يمينا ويمكن إيضاح ذلك بيانيا فيما يلى : -





الفصل الثالث

تشتت البيانات الإحصائية

أولا : تمهيد :

تبين لنا مما سبق دراسته فى الأبواب السابق إمكان تكوين وصف لبيانات المتغير الإحصائي محل الدراسة سواء بتلخيصها فى شكل جدول أو رسم بيانى ، أو بتلخيصها فى رقم حسابى هو أحد مقاييس التوسط .

وبالرغم من كل ذلك فما زال هذا الوصف للبيانات غير كامل ، إذ ما زال يوجد قصور فى وصف البيانات لم يستطع الوصف السابق الكشف عنها ، مما يقودنا إلى حكم مضلل على سلوك المتغير الإحصائي محل الدراسة ، والمثال التالى يبين ذلك .

مثال

نفرض أن لدينا مجموعتين من الطلبة تم امتحانهم فى مادة الإحصاء وكانت درجات المجموعة الأولى هى{ ٦٠ ، ٦ ، ٦٥ ، ٦٣ } ، درجات المجموعة الثانية هى { ٥٠ ، ٩٠ ، ٨٠ ، ٦٠ } والمطلوب وصف هذين المتغيرين (أو هاتين المجموعتين) بهدف المقارنة بينهما.



الحـــل

١ – نستخدم فى الوصف المتوسط الحسابى لأنه الأفضل .

∴ س للمجموعة الأولى = $\frac{٦٣ + ٦٥ + ٦٢ + ٦٠}{٤}$ = ٦٢,٥

، س للمجموعة الثانيــة = $\frac{٦٠ + ٨٠ + ٩٠ + ٢٠}{٤}$ = ٦٢,٥

٢ – ويتضح من ذلك أن متوسطة المجموعتين متساويين مما يقودنا إلى الحكم بان المجموعتين من الطلبة ذا مستوى واحـــد فـى درجات الامتحان ، فى حين أن المجموعة الأولى متقاربــة ولا يوجد بها طالب واحد راسب ، وان المجموعة الثانية متبــاعدة وبها طالب راسب ، وبالتالى فالحكم السابق حكما مضللا .

وعلى ذلك فلا بد من كشف هذا التباعد والتقارب حتى لا يقودنــا الاستنتاج إلى حكم مضلل .

والمقاييس المستخدمة فى قياس هذا التباعد والتقــارب تسـمى بمقــاييس التشتت Dispersion وهى : –

١– المدى Range
٢– الانحراف الربيعي Quartile deviation
٣– الانحراف المتوسط Mean deviation
٤– التباين Variance
٥– الانحراف المعيارى Standard deviation



ثانيا : مقاييس التشتت فى حالة المجتمع الإحصائي الصغير : -

مثال

إذا كانت مجموعة المفردات لمتغير إحصائي ما هى { ٦٠ ، ٤٠ ، ٣٠، ٧٠ ، ٥٠ ، ٩٠ ، ١٠ } فاحسب التشتت لمفردات هذه المجموعة باستخدام مقاييس التشتت المختلفة .

الحـــل

١- المـــدى : -

يعرف المدى بأنه الفرق بين اكبر قيمة واصغر قيمة فى مجموعة مفردات المتغير الإحصائي محل الدراسة .

∴ د = ى - أ

حيث : -

د : هى قيمة المدى

، ى : هى اكبر قيمة فى المجموعة

، أ : هى اصغر قيمة فى المجموعة

والاساس فى حساب المدى هو الاعتماد على القيميتين المتطرفتين فى المجموعة.

∴ د = ٩٠ - ١٠

= ٨٠

٢- الانحراف الربيعي : -

يعرف الانحراف الربيعى بأنه نصف المدى بين قيمة الربيع الاعلى وقيمة الربيع الادنى للبيانات .

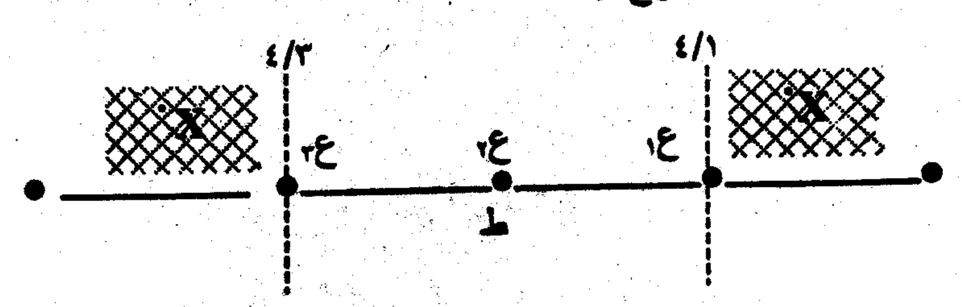
$$.. \text{الانحراف الربيع} = \frac{ع_3 - ع_1}{2}$$

حيث :

$ع_3$: هى قيمة المفردة التى ترتيبها يقع عند ¾ عدد البيانات المرتبة ترتيبا تصاعديا أو تنازليا .

$ع_1$: هى قيمة المفردة التى ترتيبها يقع عند ¼ عدد البيانات المرتبة ترتيبا تصاعديا أو تنازليا .

$ع_3 - ع_1$: تعنى انه قد تم استبعاد الربع الادنى وأيضا الربع الاعلى من البيانات ، وهذا العرض البيانى يزيد الشرح فهما .



ويرجع الاساس فى تركيب مقياس الانحراف الربيعي إلى ضرورة تلافى عيوب القياس السابق عليه (المدى) ، حيث انه عند حساب المدى كان يتم إهمال كل القيم لمفردات المجموعة ما عدا القيمتين

* العلامة × تعنى استبعاد القيم .



المتطرفتين ، وفى هذا تضليل فى حساب التشتيت فى كثير من الحالات ، وسوف يتم إيضاح هذا التضليل عند دراسة خصائص مقاييس التشتت فى نهاية هذا الفصل .

ولايجاد قيمة الانحراف الربيعى لمجموعة البيانات التـــى نحـــن بصددهـــا وهـــى { ٦٠ ، ٤٠ ، ٣٠ ، ٧٠ ، ٥٠ ، ٩٠ ، ١٠ } نتبـــع الاتـــــى :-

- ترتيب قيم مفردات المجموعة ترتيبا تنازليا (مثلا) إذ يمكـــن أن يكون الترتيب تصاعديا١٠ ، ٣٠ ، ٤٠ ، ٥٠ ، ٦٠ ، ٧٠ ، ٩٠

.. قيمة $ع_١$

.. رتبة $ع_١$ = $\frac{١}{٤}$ ن أى ربع عدد القيم

= $\frac{١}{٤}$ × ٧ = ١,٧٥ .

.. قيمة $ع_١$ هى قيمة المفردة التى ترتيبها ١,٧٥

، .. قيمة المفردة التى ترتيبها ١,٧٥ غير ملحوظ على الـــترتيب بشكل مباشر .

.. قيمة المفردة التى ترتيبها ١,٧٥ لابد وان تقع فى الفترة بين قيمـــة المفردة الأولى وقيمة المفردة الثانية علـــى الـــترتيب أي (١٠ ، ٣٠) ، وتسمى هذه الفتره بالفترة الترتيبية .



$$\therefore \text{قيمة } ع_{١} = ب + \frac{ت_{١} \times ل}{ت_{١} + ت_{٢}}$$ [١٩، ٢٠]

$$= ١٠ + \frac{(١,٧٥ - ١) \times (٣٠-١٠)}{٠,٧٥ + (٢ - ١,٧٥)}$$

$$= ١٠ + \frac{١٥}{١}$$

$$= ٢٥$$

– بالمثل توجد قيمة $ع_{٣}$

$$\therefore ع_{٣} = ٦٢,٥$$

$$\therefore \text{الانحراف الربيعى} = \frac{ع_{٣} \times ع_{١}}{٢}$$

$$= \frac{٦٢,٥ - ٢٥}{٢}$$

$$= ١٨,٧٥$$

٢- الانحراف المتوسط : -

جئ بهذا المقياس ليعالج عيوب المقاييس السابقة عليه ، فقد لاحظنا فى المدى انه يتم إهمال كل المفردات ما عدا مفردتين ، وفى

---

[١] حيث ب : هى قيمة بداية الفترة الترتيبية

$ت_{١}$: الفرق بين رتبة $ع_{١}$ والرتبة السابقة لها

$ت_{٢}$: الفرق بين $ع_{١}$ والرتبة اللاحقة لها

ل : طول الفترة الترتيبية (الفرق بين قيمة الرتبة اللاحقة والسابقة

[٢٠] تم التوصل إلى هذه المعادلة باستخدام التناسب فى الجبر



الانحراف الربيعي يتم إهمال نصف المفردات ، وفى هـذا وذلك يكـون استنتاج الحكم استنتاجا مضللا .

فالانحراف المتوسط جاء ليأخذ كل المفردات فى الاعتبار علـى أساس انه متوسط مجموع انحرافات القيم عن متوسطها الحسابى بصـرف النظر عن الإشارات * .

$$\therefore \text{الانحراف المتوسط} = \frac{\text{مجـ} (س - \bar{س})}{ن}$$

حيث :-

س : هى قيم المفردات $س_1$ ، $س_2$ ، .... ، $س_ن$

، $\bar{س}$: هى $\frac{\text{مجـ}\, س}{ن}$

ولايجاد قيمة الانحراف المتوسط لمجموعة البيانات التى نحن بصددها وهى { ٦٠ ، ٤٠ ، ٣٠ ، ٧٠ ، ٥٠ ، ٩٠ ، ١٠ } نتبع الاتى : -

$$\therefore \bar{س} = \frac{\text{مجـ}\, س}{ن}$$

$$\therefore \bar{س} = \frac{٣٥٠}{٧} = ٥٠$$

$$، \therefore \text{الانحراف المتوسط} = \frac{\text{مجـ} (س - \bar{س})}{ن}$$

$$\therefore \text{الانحراف المتوسط} = \frac{\text{مجـ} \left[(٦٠-٥٠) + (٤٠-٥٠) + (٣٠-٥٠) + (٧٠-٥٠) + (٥٠-٥٠) + (٩٠-٥٠) + (١٠-٥٠) \right]}{٧}$$

---

* لماذا ؟



$$= \frac{١٠ + ١٠ + ٢٠ + ٢٠ + ٠ + ٤٠ + ٤٠}{٧}$$

$$= \frac{١٤٠}{٧}$$

$$= ٢٠$$

٤- التباين : -

جئ بهذا المقياس ليعالج عيوب المقياس السابق عليه وهى إهمال الاشارات وذلك على أساس أن التباين هو مجموع مربعات انحرافات القيم عن متوسطها الحسابى .

$$\therefore ع^{٢} = \frac{مجـ (س - \bar{س})^{٢}}{ن}$$

حيث :-

$ع^{٢}$: هى قيمة التباين

، س ، $\bar{س}$: سبق تعريف المقصود بهما .

ولإيجاد قيمة التباين لمجموعة البيانات التى نحن بصددها وهى { ٦٠ ، ٤٠ ، ٧٠ ، ٣٠ ، ٥٠ ، ٩٠ ، ١٠ } نتبع الاتى : -

$$\therefore \bar{س} = \frac{مجـ س}{ن}$$

$$\therefore \bar{س} = \frac{٣٥٠}{٧}$$

$$= ٥٠$$



$$\therefore \text{ع}^2 = \frac{\text{مجـ}(\text{س} - \bar{\text{س}})^2}{\text{ن}}$$

$$\therefore \text{ع}^2 = \frac{[(٦٠-٥٠)^2 + (٤٠-٥٠)^2 + (٧٠-٥٠)^2 + (٣٠-٥٠)^2 + (٥٠-٥٠)^2 + (٩٠-٥٠)^2 + (١٠-٥٠)^2]}{٧}$$

$$= \frac{[(١٠)^2 + (-١٠)^2 + (٢٠)^2 + (-٢٠)^2 + (٠)^2 + (٤٠)^2 + (-٤٠)^2]}{٧}$$

$$= \frac{١٠٠ + ١٠٠ + ٤٠٠ + ٤٠٠ + ٠ + ٤٠٠ + ٤٠٠}{٧}$$

$$= \frac{١٨٠٠}{٧}$$

$$= ٢٥٧,١٤$$

٥- الانحراف المعياري : -

جئ بهذا المقياس ليعالج عيوب المقياس السابق عليه وهى تربيع الانحرافات وذلك على أساس أن الانحراف المعيارى هو الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن متوسطها الحسابى .

$$\therefore \text{ع} = \sqrt{\frac{\text{مجـ}(\text{س} - \bar{\text{س}})^2}{\text{ن}}}$$

حيث : -

ع : هى قيمة الانحراف المعيارى



، س ، سَ : سبق تعريف المقصود بهما .

ولإيجاد قيمة الانحراف المعياري لمجموعة البيانات التى نحن بصددها وهى { ٦٠ ، ٤٠ ، ٧٠ ، ٣٠ ، ٥٠ ، ٩٠ ، ١٠ } نتبع الاتى : -

$$ع = \sqrt{\frac{١٨٠٠}{٧}}$$

$$= \sqrt{٢٥٧,١٤}$$

$$= ١٦,٠٣٦$$

ثالثا : مقاييس التشتت فى حالة المجتمع الإحصائي الكبير : -

مثال

التوزيع التكرارى التالى هو توزيع درجات امتحان ٢٠ طالب فى مادة الإحصاء :

الفئات	(٠-١٠)	(١٠-٢٠)	(٢٠-٣٠)	(٣٠-٤٠)	المجموع
التكرارات	٤	٩	٥	٢	٢٠

والمطلوب :-

احسب التشتت بين درجات امتحان الـ ٢٠ طالب فى مادة الإحصاء باستخدام مقاييس التشتت المختلفة .



الحــــل

١- المــــــدى :-

.. د = ى - أ*

.. د = ٤٠ - ٠

= ٤٠

٢- الانحراف الربيعي :-

$$\text{.. الانحراف الربيعى} = \frac{ع_٣ \times ع_١}{٢}$$

.. لايجاد قيمة $ع_٣$ ، $ع_١$ نتبع الاتى :-

تكوين الجدول الإحصائي اللازم :-

الفئات	التكرار(ك)	مركز الفئة (س)	لوس
		الحدود العليا للفئات	تكرار متجمع صاعد
(٠ - ١٠)	٤	-١٠	٤
(١٠ - ٢٠)	٩	-٢٠	١٣
(٢٠ - ٣٠)	٥	-٣٠	١٨
(٣٠ - ٤٠)	٢	-٤٠	٢٠
المجموع	٢٠	/	/

---

*حيث ى : هى نهاية اكبر فئة فى الجدول

أ : هو بداية اصغر فئة فى الجدول



، ∴ $ع_٣ = \frac{٣}{٤}$ ن ∴ رتبة $ع_٣ = \frac{٣}{٤} × ٢٠ = ١٥$

∴ رتبة $ع_٣$ تقع بين ١٣ ، ١٨ من على عمود تكرار المجتمع الصاعد

$$∴ ع_٣ = ب + \frac{ت_١ × ل}{ت_١ + ت_٢}$$

$$= ٢٠ + \frac{٢ × ١٠}{٢ + ٣}$$

$$= ٢٠ + ٤$$

$$= ٢٤$$

، ∴ رتبة $ع_١ = \frac{١}{٤}$ ن ∴ $ع_١ = \frac{١}{٤} × ٢٠ = ٥$

∴ رتبة $ع_١$ تقع بين ٤ ، ١٣ من على عمود تكرار المجتمع الصاعد

$$∴ قيمة ع_١ = ب + \frac{ت_١ × ل}{ت_١ + ت_٢}$$

$$= ١٠ + \frac{١ × ١٠}{١ + ٨}$$

$$= ١٠ + \frac{١٠}{٩}$$

$$= ١١,١١$$

---

* $ت_١$ ، $ت_٢$ هى عبارة عن فرق رتب



.. الانحراف الربيعى = $\frac{24 - 11,11}{2}$

= ٦,٤٤٥

ويمكن إيضاح رتبة وقيمة الربيع الادنى والربيع الاعلى بيانيا كما يلي: -

١.رسم المنحنى المتجمع الصاعد أو الهابط أو هما معا

٢- رتبة $ع_1$ والتى تساوى ٥ على المحور الرأسي قيمتها تساوى ١١,١١ من على المحور الأفقي .

٣- رتبة $ع_3$ والتى تساوى ١٥ على المحور الرأسي قيمتها تساوى ٢٤ من على المحور الأفقي .

٤- رتبة ط (الوسيط) والتى تساوى ١٠ على المحور الرأسى قيمتها تساوى ١٦,٧ من على المحور الأفقي .

٣- لانحراف المتوسط : -

.. الانحراف المتوسط = $\frac{\text{مج (س - سَ) ك}^{*}}{\text{مج ك}}$ بصرف النظر عن الاشارات

.. لايجاد قيمة الانحراف المتوسط نتبع الاتى : -

تكوين الجدول الإحصائي اللازم : -

الفئات	تكرارات(ك)	س	ك س	سَ	(س -سَ)**	(س -سَ)ك
(٠ - ١٠)	٤	٥	٢٠		- ١٢,٥	٥٠
(١٠ - ٢٠)	٩	١٥	١٣٥		- ٢,٥	٢٢,٥
(٢٠ - ٣٠)	٥	٢٥	١٢٥		- ٧,٥	٣٧,٥
(٣٠ - ٤٠)	٢	٣٥	٧٠		- ١٧,٥	٣٥
المجموع	٢٠	-	٣٥٠	١٧,٥	-	١٤٥

.. الانحراف المتوسط = $\frac{١٤٥}{٢٠}$ = ٧,٢٥

• تم الترجيح بـــ ك أي بعدد التكرارات فى القانون (الانحراف المتوسط = $\frac{\text{مج (س-س)}}{\text{ن}}$)

** من خصائص المتوسط الحسابى أن مج (س-س) جبريا = صفر لكن مج (س-س) جبريـــا فى هذه الحالة لن يساوى الصفر فهو يساوى (+١٠) والسبب أن س فى هذه الحالة هـــى مراكز الفئات وليست المفردات الأصلية



٤- التبـــــاين : -

$$\therefore ع^٢ = \frac{مج (س - \bar{س})^٢ ك}{ك}$$

∴ لإيجاد قيمة $ع^٢$ نتبع الآتى : -

تكوين الجدول الإحصائي اللازم : -

الفئات	تكرارات(ك)	س	ك س	$\bar{س}$	$(س - \bar{س})$	$(س - \bar{س})^٢$	$(س - \bar{س})^٢ ك$
(٠ - ١٠)	٤	٥	٢٠		- ١٢,٥	١٥٦,٢٥	- ٦٢٥,٠
(١٠ - ٢٠)	٩	١٥	١٣٥		- ٢,٥	٦,٢٥	- ٥٦,٢٥
(٢٠ - ٣٠)	٥	٢٥	١٢٥		- ٧,٥	٥٦,٢٥	- ٢٨١,٢٥
(٣٠ - ٤٠)	٢	٣٥	٧٠		- ١٧,٥	٣٠٦,٢٥	- ٦١٢,٥٠
المجموع	٢٠	-	٣٥٠	١٧,٥	-		١٥٧٥

$$\therefore ع^٢ = \frac{١٥٧٥}{٢٠} = ٧٨,٧٥$$

ملحـــــوظة : -

يلاحظ في العمليات الحسابية وجود كسر عشرى من رقمين فى قيمة الوسط الحسابى ، وفى كثير من الاحيان يكون الكسر العشرى هذا مكون من ثلاثة أرقام أو اربعة أو .... ، مما نضطر معه إجراء عملية التقريب وبالتالى تكون الانحرافات الناتجة (س - $\bar{س}$) مقربة وليست بـالضبط فينتج عن ذلك خطأ تراكمى فى قيمة التباين .

---

*تم الترجيح بـ ك أي بعدد التكرارات فى القانون $\left(ع^٢ = \frac{مج (س - \bar{س})^٢}{ن} \right)$

وللتخلص من هذه المشكلة أمكن التوصل إلى الطريقة التالية لقانون التباين : -

$$\therefore \text{مجـ}(\text{س} - \bar{\text{س}})^{٢} = \text{مجـ}(\text{س}^{٢} - ٢\,\text{س}\,\bar{\text{س}} + \bar{\text{س}}^{٢}) \quad \text{.... جبريا}$$

$$= \text{مجـ س}^{٢} - ٢\,\text{مجـ س}\,\bar{\text{س}} + \text{ن}\,\bar{\text{س}}^{٢} \quad \text{... جبريا واحصائيا}^{(1)}$$

$$= \text{مجـ س}^{٢} - ٢\,\text{مجـ س} \times \frac{\text{مجـ س}}{\text{ن}} + \text{ن}\left(\frac{\text{مجـ س}}{\text{ن}}\right)^{٢} \quad \text{جبريا ، واحصائيا}^{(2)}$$

$$= \text{مجـ س}^{٢} - ٢\,\frac{(\text{مجـ س})^{٢}}{\text{ن}} + \frac{(\text{مجـ س})^{٢}}{\text{ن}}$$

$$= \text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{\text{ن}}$$

$$\text{،} \therefore \text{ع}^{٢} = \frac{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}}{\text{ن}}$$

$$\therefore \text{ع}^{٢} = \frac{\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{\text{ن}}}{\text{ن}}$$

، وفى حالة المجتمع الإحصائي الكبير (البيانات المبوبة)

$$- \quad \text{ع}^{٢} = \frac{\text{مجـ س ك}^{٢} - \frac{(\text{مجـ س ك})^{٢}}{\text{ن}}}{\text{مجـ ك}}$$

(1) مجـ س = ن $\bar{\text{س}}$ لان المجموع فى هذه الحالة هو جمع متكرر لرقم ثابت

(2) $\text{س} = \frac{\text{مجـ}\,\bar{\text{س}}}{\text{ن}}$

* هذا القانون يتعامل مع مربعات القيم مباشرة دون ما نحتاج لإيجاد الوسط الحسابى ثم الانحرافات ثم ......



وباستخدام هذا القانون على بيانات المثال السابق نتبع الاتى : -

تكوين الجدول الإحصائي اللازم : -

الفئات	التكرار (ك)	مركز الفئة س	س ك	س^{٢}	س^{٢} ك
(٠ – ١٠)	٤	٥	٢٠	٢٥	١٠٠
(١٠ – ٢٠)	٩	١٥	١٣٥	٢٢٥	٢٠٢٥
(٢٠ – ٣٠)	٥	٢٥	١٢٥	٦٢٥	٣١٢٥
(٣٠ – ٤٠)	٢	٣٥	٧٠	١٢٢٥	٢٤٥٠
المجموع	٢٠	-	٣٥٠	٢١٠٠	٧٧٠٠

$$\therefore ع^{٢} = \frac{٧٧٠٠ - \frac{(٣٥٠)^{٢}}{٢٠}}{٢٠}$$

$$= \frac{٧٧٠٠ - ٦١٢٥}{٢٠}$$

$$= ٧٨,٧٥$$ وهى نفس النتيجة السابقة

٥- الانحراف المعيارى : -

$$\therefore ع = \sqrt{\frac{مج (س - \bar{س})^{٢}}{ن}}$$

$$أو\ ع = \sqrt{\frac{مج س^{٢} ك - \frac{(مج س ك)^{٢}}{مج ك}}{مج ك}}$$



$$\therefore ع = \sqrt{٧٨,٧٥}$$

$$= ٨,٨٧$$

خواص مقاييس التشتت : -

سبق أن أوضحنا أن مقاييس التشتت كلها معيبة سواء التى تعتمد فى القياس على قيمتين أو التى تعتمد على قيمة واحدة ، إلا أن الانحراف المعيارى يعتبر ادق مقاييس التشتت لانه يعتمد فى القياس على حساب انحرافات القيم عن متوسطها الحسابى الذى يعتبر مركز الثقل للمجموعة ومن ثم فالانحراف المعيارى يمثل الحد الادنى ويمكن اثبات ذلك رياضيا فيما يلى : -

.. المطلوب هو اثبات أن متوسط مجموع مربعات انحرافات القيم عن متوسطها الحسابى اقل ما يمكن أي اقل من متوسط مجموع مربعات انحرافات القيم عن أي قيمة أخرى غير المتوسط الحسابى . ، نفرض اننا رمزنا للوسط الحسابى بالرمز سَ وللقيمة الأخرى بالرمز هـ

.. المطلوب هو اثبات أن $\frac{مج (س - هـ)^٢}{ن} > \frac{مج (س - \bar{س})^٢}{ن}$

، .. س - هـ = س - سَ + سَ - هـ (خواص المتساوية)

$$\therefore مج (س - هـ)^٢ = مج (س - \bar{س} + \bar{س} - هـ)^٢$$



= مجـ [(س - سَ) + (سَ - هـ)]٢

= مجـ [(س - سَ)٢ + ٢(س - سَ) (سَ - هـ) + (سَ - هـ)]٢

= مجـ (س - سَ)٢ + ٢ مجـ (س - سَ) (سَ - هـ) + مجـ (سَ - هـ)٢

= مجـ (س - سَ)٢ + مجـ (سَ - هـ)٢ حيث مجـ (س - سَ)

= صفر

∴ الطرف الايمن اكبر من أي من حدى الطرف الايسر

∴ مجـ (س - هـ)٢ > مجـ (س - سَ)٢

هـ . ط . ث



موضوعات هامة فى التشتت عند التطبيق : -

١- تصحيح شبرد للتباين : -

كنا نفترض عند حساب المتوسط الحسابى من التوزيع التكرارى أن مركز الفئة يمثل كل مفردات الفئة فمثلاً الفئة (٠ – ١٠) مركزها ٥ ومن ثم فالقيمة ٥ ليست هى المفردة ١ أو ٢ أو ٣ ... أو ١٠ وانما هى ممثلة لتلك المفردات ، وبناء على هذا الفرض تتم الحسابات فى إيجاد المتوسط الحسابى ، وهنا يتضح وجود شئ من الخطأ فى هذا الحساب عما لو تم على أساس كل المفردات .

وقد قام شبرد بتصحيح هذا الخطأ وذلك بطرح $\frac{أ^٢}{١٢}$ (حيث أ هى طول الفئة) من قيمة التباين ثم اخذ الجذر التربيعي لنتيجة الطرح فينتج الانحراف المعيارى الصحيح.

مثال

إذا كان التباين لاحد التوزيعات التكرارية هو ١١١ ، وطول الفئة ١٠ فاوجد الانحراف المعيارى المصحح .

الحل

.. الانحراف المعيارى المصحح = $\sqrt{\text{التباين} - \frac{أ^٢}{١٢}}$

$= \sqrt{١١١ - \frac{١٠^٢}{١٢}}$

$= ١٠,١١$

٢- التشتت النسبي :-

عند المقارنة بين تشتت مجموعتين مختلفتين فى وحدات القياس أو فى وسطهما الحسابى ، فانه يجب استخدام معامل الاختلاف عند إجراء عملية المقارنة وذلك للتخلص من هذين الاختلافين .

مثال (١)

إذا كان متوسط الطول لمجموعة من الأفراد هو ١٥٠ سم بانحراف معيارى ١٠ سم، وكان متوسط العمر لافراد هذه المجموعة هو ٤٠ سنة بانحراف معيارى ٣ سنوات ، فاذكر أي المجموعتين اكثر تشتتا من الأخرى .

الحل

من الخطأ القول أن المجموعة الأولى اكثر تشتتا من المجموعة الثانية لمجرد ملاحظة أن الانحراف المعيارى للاولى ١٠ وللثانية ٢ ذلك لان الرقم ١٠ يعبر عن السنتيمترات بينما الرقم ٢ يعبر عن السنوات . والمقارنة الصحيحة هى التى تستخدم التشتت النسبي حيث النسبة لا تميز، ويقاس التشتت النسبي بمعامل الاختلاف Coefficient of variation.

$$\text{معامل الاختلاف} = \frac{ع}{\bar{س}} \times ١٠٠$$

حيث :

ع : هى الانحراف المعيارى

$\bar{س}$: هى المتوسط الحسابى



.. معامل الاختلاف للمجموعة الأولى = $\frac{١٠}{١٥٠} \times ١٠٠ = ٦,٧\%$

و معامل الاختلاف للمجموعة الثانية = $\frac{٣}{٤٠} \times ١٠٠ = ٧,٥\%$

اى أن المجموعة الثانية هى الأكثر تشتتا من المجموعة الاولى

مثال (٢)

إذا كان متوسط الإنفاق اليومى لمجموعة من السائحين الالمان على الطعام هو ٥ دولار بانحراف معيارى ٠,٢ دولار ، بينما متوسط الإنفاق اليومى لمجموعة من السائحين العرب على الطعام هو ٥,٢٥ دولار بانحراف معيارى٠,٣ دولار ، فاذكر أي المجموعتين اكثر تشتتا .

الحل

معامل الاختلاف للمجموعة الأولى = $\frac{٠,٢}{٥} \times ١٠٠ = ٠,٤\%$

معامل الاختلاف للمجموعة الثانية = $\frac{٠,٣}{٥,٢٥} \times ١٠٠ = ٥,٧\%$

أي أن المجموعة الثانية هى الأكثر تشتتا من المجموعة الأولى

٣- الدرجة المعيارية :-

عند المقارنة بين مفردتين تنتميان إلى مجموعتين مختلفتين فانه يجب إيجاد الترتيب النسبي للمفردة داخل مجموعتها ، ويتحقق ذلك بتحويل الفرق بين قيمة المفردة ومتوسطها الحسابى إلى فرق معيارى ، وهذا ما يسمى بالدرجة المعيارية Standardized score.

$$ى = \frac{س - \bar{س}}{ع}$$

حيث:

ى : هى الدرجة المعيارية البديلة عن قيمة المفردة س .

س : هى قيمة المفردة فى المجموعة .

$\bar{س}$: المتوسط الحسابى للمجموعة .

ع : الانحراف المعيارى للمجموعة .

مثال

كان المتوسط الحسابى لدرجات الطلاب فى مادة الإحصاء هو ٧٠ درجة بانحراف معيارى ٥ درجات ، وكان المتوسط الحسابى لدرجات الطلاب فى مادة الاقتصاد هو ٦٥ درجة بانحراف معيارى ٢ درجة ، وكان أحد الطلاب قد حصل على ٨٠ درجة فى الإحصاء ، ٧٠ درجة فى الاقتصاد ، فقارن بين أداء الطالب فى المادتين .

الحل

من الخطأ القول أن أداء هذا الطالب فى مادة الإحصاء افضل من أداؤه فى مادة الاقتصاد لمجرد ملاحظة الدرجتين ٨٠ ، ٧٠ إذ يجب أن يؤخذ فى الاعتبار توزيع الدرجات فى المادتين حتى تكون المقارنة سليمة وهذا ما يدعو إلى استخدام الدرجة المعيارية .

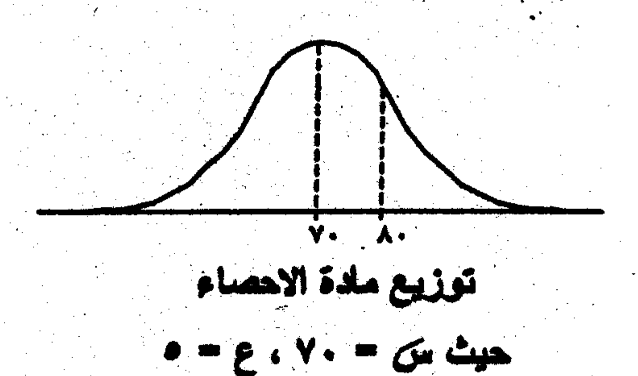
$$\therefore ى = \frac{س - \bar{س}}{ع}$$

$\therefore$ الدرجة الاحصائية للطالب = $\frac{٨٠ - ٧٠}{٥}$ = ٢ درجة معيارية

، ى لدرجة الاقتصاد للطالب = $\frac{٧٠ - ٦٥}{٢}$ = ٢,٥ درجة معيارية

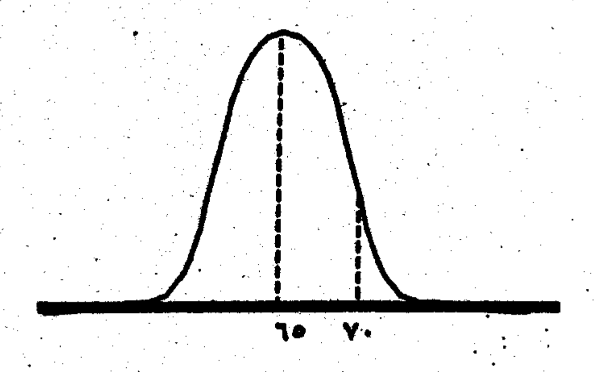
من ذلك نستنتج أن أداء الطالب فى مادة الاقتصاد افضل من أداؤه فى مادة الإحصاء.

ويمكن إيضاح ذلك بيانيا كما يلى :-



توزيع مادة الاحصاء

حيث $\bar{س}$ = ٧٠ ، ع = ٥



توزيع مادة الاقتصاد

حيث $\bar{س}$ = ٦٥ ، ع = ٢



٤- استخدام مقاييس التوسط والتشتت معا فى وصف التوزيع -

سبق استخدام مقاييس التوسط في وصف التواء التوزيع حيث بيان نوع الالتواء يمينا أم يسارا لكن دون قياس درجة الالتواء هل شديد أم ضئيل ، وانه باستخدام مقاييس التوسط والتشتت معا يمكن وصف التواء التوزيع من حيث النوع والدرجة : -

- اذا كان $\frac{\bar{\text{س}} - \text{ط}}{\text{ع}} = ٠$ كان التوزيع متماثل

- اذا كان $\frac{\bar{\text{س}} - \text{ط}}{\text{ع}} = -٠,٢$ كان التوزيع ملتوى يسار ضئيل

- اذا كان $\frac{\bar{\text{س}} - \text{ط}}{\text{ع}} = +٠,٨$ كان التوزيع ملتوى يمينا وبشده

.. معامل الالتواء = $\frac{\bar{\text{س}} - \text{ط}}{\text{ع}}$ تتراوح قيمته بين (-١ ، +١)

ويمكن أن يكون معامل الالتواء = $\frac{\bar{\text{س}} - \text{م}}{\text{ع}}$

لكن عند مقارنة درجة الالتواء لعدة توزيعات مختلفة يشترط استخدام نفس المعامل .

العزوم فى الإحصاء واهميتها فى وصف التوزيع : -

يعتبر ادق وصف لالتواء وتفرطح التوزيع ذلك الذى يقوم على استخدام العزوم ، وفيما يلى فكرة عن العزم فى الإحصاء : -

العزم الأول حول الصفر = $\frac{\text{مجـ}(\text{س} - ٠)\text{ك}}{\text{مجـ ك}} = \frac{\text{مجـ س ك}}{\text{مجـ ك}} = \bar{\text{س}}$

العزم الثانى حول الصفر = $\frac{\text{مجـ}(\text{س} - ٠)^{٢}\text{ك}}{\text{مجـ ك}} = \frac{\text{مجـ س}^{٢}\text{ ك}}{\text{مجـ ك}}$



العزم الثالث حول الصفر = $\frac{مج (س - ٠)^٣ ك}{مج ك} = \frac{مج س^٣ ك}{مج ك}$

-

-

-

-

العزم ن حول الصفــر = $\frac{مج (س - ٠)^ن ك}{مج ك} = \frac{مج س^ن ك}{مج ك}$

وبالمثل :

العزم الأول حول الوسط الحسابى = $\frac{مج (س - \bar{س}) ك}{مج ك}$ = صفر

العزم الثانى حول الوسط الحسابى = $\frac{مج (س - \bar{س})^٢ ك}{مج ك}$ = التباين

العزم الثالث حول الوسط الحسابى = $\frac{مج (س - \bar{س})^٣ ك}{مج ك}$

العزم الرابع حول الوسط الحسابى = $\frac{مج (س - \bar{س})^٤ ك}{مج ك}$

-

-

-

وهكذا

معامل الالتواء = $\frac{م_٣^٢}{م_٢^٣}$ = $\frac{\text{مربع العزم الثالث}}{\text{مكعب العزم الثانى}}$



ملحوظة : -

إذا ذكر العزم فقط فالمقصود هو العزم حول الوسط الحسابى

$$\text{معامل التفرطح} = \frac{م_4}{م_2^2} = \frac{\text{العزم الرابع}}{\text{مربع العزم الثانى}} = \frac{\text{العزم الرابع}}{\text{مربع التباين}}$$

ولما كان معامل تفرطح التوزيع الطبيعى هو ٣ فانه يتم مقارنة معامل التفرطح لاى توزيع بالرقم ٣ ، فان ذاد عن ٣ كان التوزيع مدبب وان قل عن ٣ كان التوزيع مفرطح .

مثال

فيما يلى بيانات عن توزيع عدد من العمال حسب الأجر الشهرى:-

الأجر الشهرى	(١٠-١٤)	(١٤-١٨)	(١٨-٢٢)	(٢٢-٢٦)	(٢٦-٣٠)	(٣٠-٣٤)	(٣٤-٣٦)	المجموع
عدد العمال	١٠٠	٣٢٠	٢٥٢	٢٠٨	٦٠	٤٤	١٦	١٠٠٠

والمطلوب : -

قياس الالتواء والتفرطح مع وصف التوزيع .



الحـــل

تكوين الجدول الإحصائي اللازم * : -

الفئات	تكرارات(ك)	مركز الفئة س	ح	ح ك	$ح^{٢}ك$	$ح^{٣}ك$	$ح^{٤}ك$
١٠ – ١٤	١٠٠	١٢	٨-	٨٠٠ -	٦٤٠٠	٥١٢٠٠-	٤٠٩٦٠٠
١٤ – ١٨	٣٢٠	١٦	٤-	١٢٨٠ -	٥١٢٠	٢٠٤٨٠-	٨١٩٢٠
١٨ – ٢٢	٣٥٣	٢٠	صفر	صفر	صفر	صفر	صفر
٢٢ – ٢٦	٢٠٨	٢٤	٤	٨٣٢	٣٣٢٨	١٣٣١٢	٥٣٢٤٨
٢٦ – ٣٠	٦٠	٢٨	٨	٤٨٠	٣٨٤٠	٣٠٧٢٠	٢٤٥٧٦٠
٣٠ – ٣٤	٤٤	٣٢	١٢	٥٢٨	٦٣٣٦	٧٦٠٣٢	٩١٢٣٨٤
٣٤ – ٣٦	١٦	٣٥	١٥	٢٤٠	٣٦٠٠	٥٤٠٠٠	٨١٠٠٠٠
المجموع	١٠٠٠	/		صفر	٢٨٦٢٤	١٠١٣٨٤	٢٥١٢٩١٢

*العزم الثانى (م٢) = $\frac{\text{مج ح}^{٢}\text{ك}}{\text{مج ك}} - \left(\frac{\text{مج ح ك}}{\text{مج ك}}\right)^{٢}$

$\therefore$ م٢ = $\frac{٢٨٦٢٤}{١٠٠٠} - \left(\frac{\text{صفر}}{١٠٠٠}\right)^{٢}$

= ٢٨,٦٢٤

، $\therefore$ العزم الثالث (م٣) = $\frac{\text{مج ح}^{٣}\text{ك}}{\text{مج ك}}$ -

$$\left[- ٣ \times \frac{\text{مج ح ك}}{\text{مج ك}} \times \frac{\text{مج ح}^{٢}\text{ك}}{\text{مج ك}} + ٢ \left(\frac{\text{مج ح ك}}{\text{مج ك}}\right)^{٢} \right]$$

---

* الجدول حتى العمود السادس هو بالضبط جدول حساب التباين أما العمودان السابع والثامن فهما لإيجاد العزم الثالث والعزم الرابع .

* هذه قوانين لها اثبات .

$$\therefore \text{م}_{٣} = \frac{١٠١٣٨٤}{١٠٠٠} - \left[٣ \times \frac{\text{صفر}}{١٠٠٠} \times \frac{٢٨٦٢٤}{١٠٠٠} + ٢ \times \left(\frac{\text{صفر}}{١٠٠٠}\right)^{٢} \right]$$

$$= ١٠١,٣٨٤ - \text{صفر} + \text{صفر}$$

$$= ١٠١,٣٨٤$$

$$\therefore \text{معامل الالتواء} = \frac{\text{م}_{٣}^{٢}}{\text{م}_{٢}^{٣}}$$

$$= \frac{(١٠١,٣٨٤)^{٢}}{(٢٨,٦٢٤)^{٣}} = ٠,٤٢$$

وهذا التواء موجب أى يتجه نحو اليمين كما أنه التواء قليل .

*العزم الرابع $(\text{م}_{٤})$ $= \frac{\text{مج ح}^{٤}\text{ك}}{\text{مج ك}} - ٣\left(\frac{\text{مج ح ك}^{٢}}{\text{مج ك}}\right)^{٤} -$

$$- \left(٤ \times \frac{\text{مج ح ك}}{\text{مج ك}} \times \frac{\text{مج ح}^{٣}\text{ك}}{\text{مج ك}} \right)$$

$$+ \left[٦ \left(\frac{\text{مج ح ك}}{\text{مج ك}}\right)^{٢} \times \frac{\text{مج ح}^{٢}\text{ك}}{\text{مج ك}} \right]$$

$$\therefore \text{م}_{٤} = \frac{٢٥١٢٩١٢}{١٠٠٠} - ٣\times\left(\frac{\text{صفر}}{١٠٠٠}\right)^{٢} - \left(٤\times \frac{\text{صفر}}{١٠٠٠} \times \frac{١٠١٣٨٤}{١٠٠٠}\right)$$

$$+ \left[٦ \left(\frac{\text{صفر}}{١٠٠٠}\right)^{٢} \times \frac{٢٨٦٢٤}{١٠٠٠} \right]$$

---

* هذا قانون له اثبات

= ٢٥١٢,٩١٢ - صفر - صفر + صفر

= ٢٥١٢,٩١٢

، .. معامل التفرطح = $\frac{م_٤}{(م_٢)^٢}$

:. معامل التفرطح = $\frac{٢٥١٢,٩١٢}{(٢٨,٦٢٤)^٢}$

= ٢,٩٧

وهذا يبين أن التوزيع قليل التفرطح حيث أن تفرطح التوزيع المعتاد ٣ .

الاثبات لأحد القوانين السابقة :-

.. العزم الثانى حول الوسط الحسابى = $\frac{مج (س - \bar{س})^٢ ك}{مج ك}$

= $\frac{مج (س^٢ - ٢ س \bar{س} + \bar{س}^٢) ك}{مج ك}$

= $\frac{مج س^٢ ك}{مج ك}$ - ٢ $\frac{مج س ك}{مج ك}$ ×

× $\frac{مج س ك}{مج ك}$ + $\left(\frac{مج س ك}{مج ك}\right)^٢$

= $\frac{مج س^٢ ك}{مج ك}$ - ٢ $\left(\frac{مج س ك}{مج ك}\right)^٢$ +

+ $\left(\frac{مج س ك}{مج ك}\right)^٢$

$$= \frac{\text{مج س}^{2}\,\text{ك}}{\text{مج ك}} - \left(\frac{\text{مج س ك}}{\text{مج ك}}\right)^{2}$$

$$\text{ومكن} = \frac{\text{مج ح}^{2}\,\text{ك}}{\text{مج ك}} - \left(\frac{\text{مج ح ك}}{\text{مج ك}}\right)^{2}$$

هـ . ط . ث

الباب الرابع

الدراسة الاحصائية للعلاقة بين متغيرين احصائيين

الدراسة الإحصائية للعلاقة بين متغيرين

يشتمل هذا الباب على فصلين : -

الفصل الأول : الارتبـــاط البسيط

- مفهوم الارتباط البسيط
- الارتباط البسيط بيانيا
- الارتباط البسيط جبريا .
- حالات فى الارتباط البسيط

الفصل الثانى : الانحـــدار البسيـــط

- مفهوم الانحدار البسيط
- الانحدار البسيط ذو الخط المستقيم :

  - $\hat{ص} = د(س) = أ + ب\,س$
  - الخطأ المعيارى للتقدير ($\hat{ص}$)
  - العلاقة بين الخطأ المعيارى للتقديـــر والمعنويـــة لمعامل الانحدار والنموذج ككل .

- الانحدار البسيط ذو الخط المنحى :

  - $\hat{ص} = د(س) = أ + ب\,س^{٢}$ (المقطع المكافىء)
  - $\hat{ص} = د(س) = أ\,س^{ب}$ (دالة هندسيـــة)
  - $\hat{ص} = د(س) = أ\,ب^{س}$ (دالة أسيـــة)



الفصل الأول

الارتباط البسيط

مقدمه :-

فى دراستنا الإحصائية التى تم تناولها فى الأبواب السابقة كنا نتعامل مع متغير احصائى واحد فى مجموعة واحدة ، وقد أمكن وصف هذا المتغير من خلال عرض بياناته سواء كان العرض جدولى أو بيانى ، ثم قياس النزعة المركزية لهذه البيانات وأيضا قياس تشتتها بالاضافة الى قياس الالتواء والتفرطح لها . لكن الأن نتعرض فى دراستنا الاحصائية للعلاقة بين متغيرين معا فى مجموعة واحدة ، كدراسة العلاقة بين الوزن والطول لمجموعة من الطلاب ، أو دراسة العلاقة بين درجات نجاح مجموعة من الطلاب فى مادتى الاحصاء والاقتصاد ، أو دراسة العلاقة بين الكمية والسعر لسلعة معينة ، أو دراسة العلاقة بين الأجر والمهنة لمجموعة من العمال ، أو دراسة العلاقة بين الدخل والاستهلاك لمجموعة من الأسر ، أو دراسة العلاقة بين جنسية السائح وانفاقه ، والعلاقة بين متغيرين (أو ظاهرتين) تسمى بالارتباط .

المقصود بالارتباط البسيط :-

يقصد بكلمة البسيط أن الارتباط فى هذه الحالة ينصب على العلاقة بين متغيرين اثنين فقط ، بمعنى أنه إذا كانت العلاقة بين اكثر من متغيرين فالارتباط البسيط لايستخدم فى هذه الحالة بل يستخدم الارتباط المتعدد و الجزئى وهذا ما سيتم التعرض له فى الباب القادم .



ويبحث الارتباط فى وجود أو عدم وجود علاقة بين المتغيرين ، وإذا وجدت هذه العلاقة فالبحث يهتم بقياس درجة هذه العلاقة أى العلاقة قوية أم ضعيفة ، كما يهتم بمعرفة نوع هذه العلاقة أى طردية (موجبة) أم عكسية (سالبة) . ويمكن إيضاح هذا المفهوم للارتباط بيانيا وجبريا .

الارتباط البسيط بيانيا :

إذا كانت العلاقة بين المتغيرين تأخذ قيمًا عددية أى أزواج مرتبة (س ، ص) فإنه يمكن عرض هذه البيانات أى النقاط فى شكل بيانى ديكارتى يسمى بشكل الانتشار **Scatter diagram** . ومن شكل الانتشار يمكن التعرف بمجرد النظر على وجود ارتباط من عدمه وأيضا تكوين فكرة معقولة عن درجة ونوع هذا الارتباط فى حالة وجوده ، والأمثلة التالية توضح ذلك .

مثال (١)

الجدول التالى يوضح بيانات عن المتغيرين س ، ص :

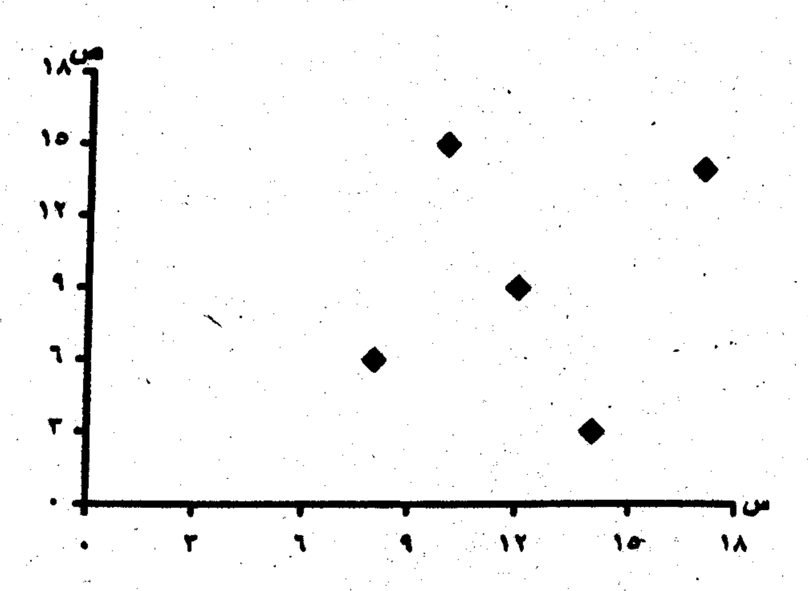
س	٨	١٢	١٠	١٧	١٤
ص	٦	٩	١٥	١٤	٣

والمطلوب :

رسم شكل الانتشار لنقاط هذا الجدول ثم استخدام شكل الانتشار الناتج فى إيضاح وجود الارتباط بين المتغيرين من عدمه .



الحــل



والشكل السابق هو شكل الانتشار ويتضح منه عدم وجود نموذج واضح لاتجاه توزيع النقط فهى تنتشر فى كل الاتجاهات ، الأمر الذى يصعب معه إيضاح خط الانتشار[1] ، وعليه لا يوجد ارتباط بين هذين المتغيرين.

مثال (٢)

الجدول التالى يوضح بيانات عن المتغيرين س ، ص :

س	١	٣	٤	٧	١٠
ص	٧	١٩	٢٥	٤٣	٦١

والمطلوب :

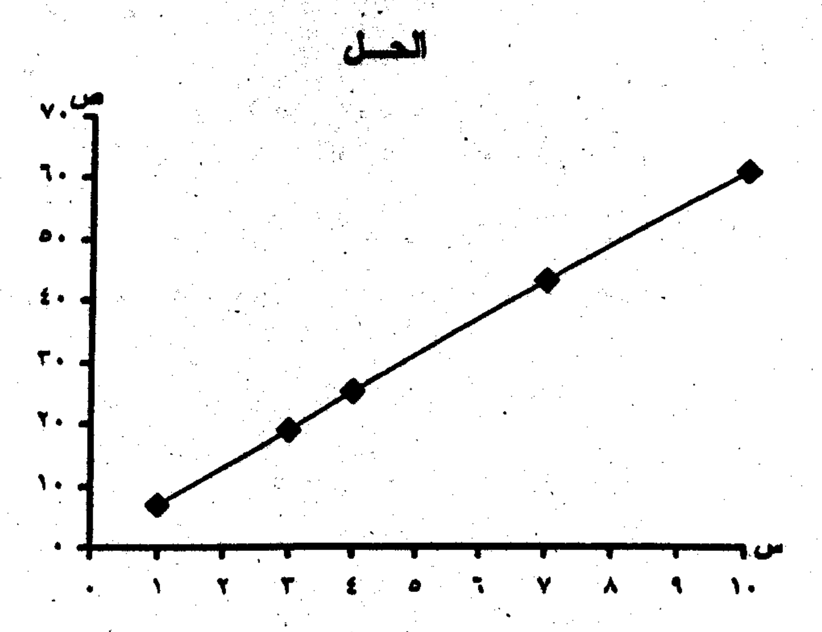
رسم شكل الانتشار لنقاط هذا الجدول ثم استخدام شكل الانتشار الناتج فى إيضاح وجود الارتباط بين المتغيرين من عدمه ، وفى حالة وجوده وضح بدرجة معقولة نوع ودرجة الارتباط .

---

(١) خط الانتشار هو خط بيانى (مستقيم أو منحنى) يتوسط اتجاه إنتشار النقط .



الحــل



الشكل السابق هو شكل الانتشار ، ويتضح منه وجــود نمــوذج واضــح لاتجاه توزيع النقط ، ولذلك يوجد خط الانتشار ، وإن كان خط الانتشـــار هو خط يتوسط اتجاه النقط إلا أنه فى هذه الحالة هو خط مستقيم تنطبــق عليه جميع النقاط ، وهنا يكون الارتباط بيــن المتغــير ارتباطــا تامــا وموجب[1] وهنا يلاحظ أن القيم الكبيرة لأحد المتغيرين يقابلها قيم كبـــيرة للمتغير الآخر والقيم الصغيرة لأحد المتغيرين يقابلها قيم صغيرة للمتغــير الآخر .

مثال (٣)

الجدول التالى يوضح بيانات عن المتغيرين س ، ص :

س	٢	٤	٨	١٠	١٦
ص	١١	١٠	٨	٧	٤

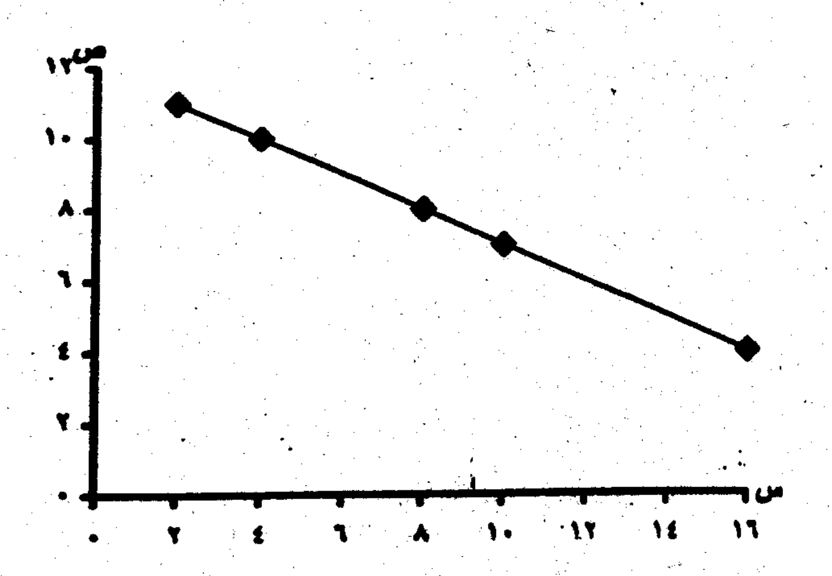
---

(١) موجب لأن ميل خط الإنتشار فى هـــذه الحالــة موجب ، والميل هندســيا هــو ظــل الزاوية التى يصنعها المستقيم مع الاتجاه الموجــب لمحــور الســينات ، ظــا الزاوية< ٩٠ ° = قيمة موجبة .

والمطلوب :

رسم شكل الانتشار لنقاط هذا الجدول ثم استخدام شكل الانتشار الناتج فى إيضاح وجود الارتباط بين المتغيرين من عدمه ، وفى حالة وجوده وضح بدرجة معقولة نوع درجة هذا الارتباط .

الحـل



الشكل الناتج هو شكل الانتشار ، ويتضح منه وجود نموذج واضح لاتجاه انتشار النقط ، ولذلك يوجد خط للانتشار ، وهو فى هذه الحالة خط مستقيم تنطبق على جميع النقط ، ولذلك فهو ارتباط تام ، كما أن نوع هذا الارتباط سالب حيث أن ميل خط الانتشار سالب لأن ظا الزاوية > ٩٠° = قيمة سالبة ، أى أن الارتباط عكسى وهنا يلاحظ أن القيم الكبيرة لأحد المتغيرين يقابلها قيم صغيرة للمتغير الأخر والقيم الصغيرة لأحد المتغيرين يقابلها قيم كبيرة للمتغير الآخر .

مثال (٤)

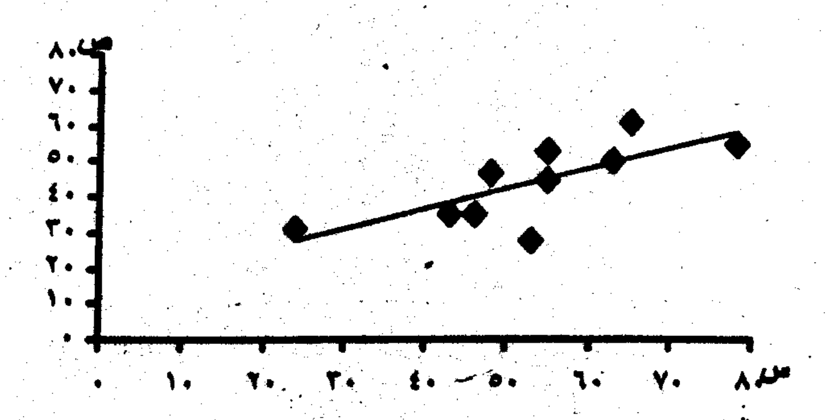
الجدول التالي يوضح بيانات عن المتغيرين س ، ص :

س	٧٨	٦٥	٦٣	٥٥	٥٣	٢٤	٤٣	٤٦	٤٨	٥٥
ص	٥٥	٦١	٥٠	٥٣	٢٨	٣١	٣٥	٣٥	٤٧	٤٥

والمطلوب :

رسم شكل الانتشار لنقاط هذا الجدول ثم استخدام شكل الانتشار الناتج فى إيضاح وجود الارتباط بين المتغيرين من عدمه ، وفى حالة وجوده وضح بدرجة معقولة نوع درجة هذا الارتباط .-

الحل



الشكل الناتج هو شكل الانتشار ، ويتضح منه وجود نموذج واضح لاتجاه انتشار النقط ، ولذلك يوجد خط للانتشار ، وهو فى هذه الحالة خط مستقيم لكن لا تنطبق عليه جميع النقط وإنما يوضح درجة انتشار النقط أى ابتعادها عنه ، فكلما ابتعدت النقط عن خط الانتشار كلما ضعفت درجة الارتباط ، وانه كلما اقتربت النقط من خط الانتشار كلما قويت درجة الارتباط ، وعليه فالارتباط فى هذه الحالة هو ارتباط قوى موجب.

مثال (٥)

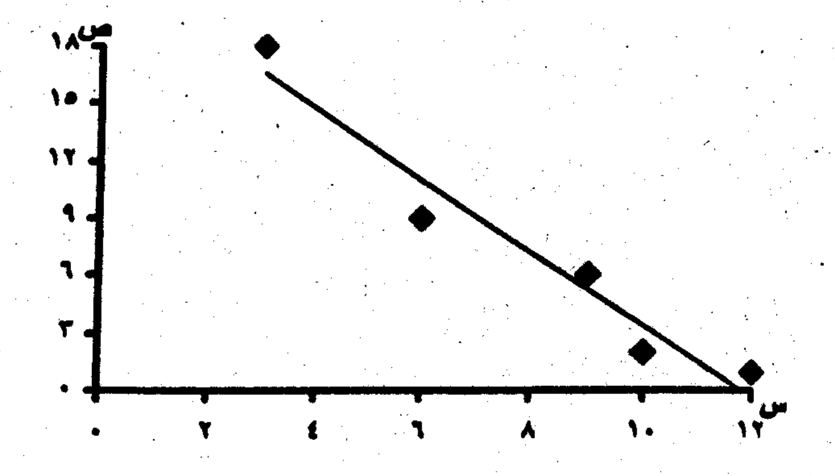
الجدول التالى يبين بيانات عن المتغيرين س ، ص :

٣	٦	٩	١٠	١٢	س
١٨	٩	٦	٢	١	ص

والمطلوب :

رسم شكل الانتشار لنقاط هذا الجدول ثم استخدام شكل الانتشار الناتج فى إيضاح وجود الارتباط من عدمه ، وفى حالة وجوده وضح بدرجة معقولة نوع ودرجة هذا الارتباط .





الشكل الناتج هو شكل الانتشار ، ويتضح منه وجود نموذج واضح لاتجاه انتشار النقط ، ولذلك يوجد خط للانتشار ، وهو فى هـذه الحالـة خـط مستقيم وعليه فالارتباط قوى سالب .

مثال (٦)

الجدول الآتى يبين بعض قيم س ، ص المناظرة لها :

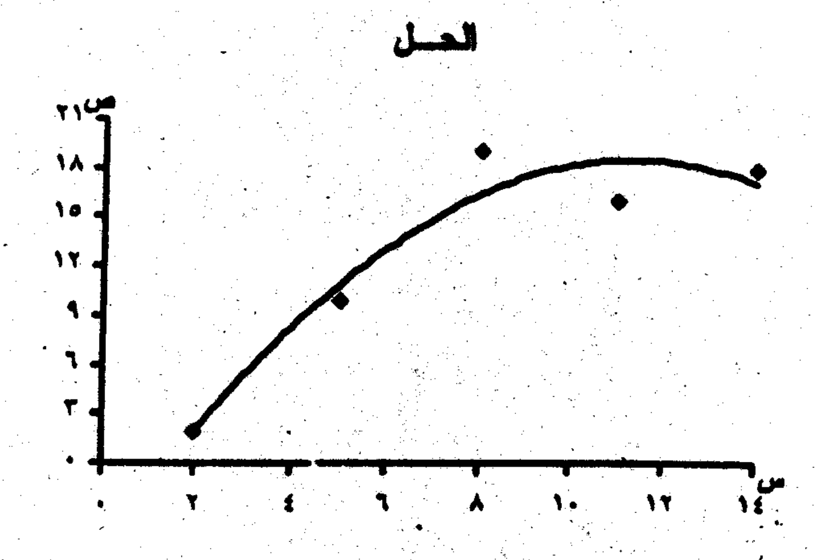
١٤	١١	٨	٥	٢	س
١٨	١٦	١٩	١٠	٢	ص



والمطلوب :

رسم شكل الانتشار لنقاط هذا الجدول واستخدامه في دراسة الارتباط بين هذين المتغيرين .

الحل



الشكل الناتج هو شكل الانتشار ، ويتضح منه وجود نموذج واضح لاتجاه إنتشار النقط ، ولذلك يوجد خط للأنتشار ، وهو في هذه الحالة خط منحني ، وعليه فالارتباط قوي(*) .

الارتباط البسيط ذو الخط المستقيم :

هو القياس الكمي للتغير الحادث في العلاقة بين متغيرين ذات قيم عددية وذلك من خلال معادلة رياضية تعرف بمعامل الارتباط البسيط ، ويعرف معامل الارتباط البسيط بأنه متوسط مجموع حاصل ضرب

(*) في هذه الحالة لا يجب القول أن العلاقة موجبة أو سالبه لأن العلاقة تكون موجبه في جزء من المنحنى وسالبة في جزء آخر ، وذلك لأن المنحنى ليس له ميل واحد وإنما له عدة ميول .



انحرافات* القيم عن متوسطها الحسابي للمتغيرين معاً والمقاس بقيم معيارية** ، والصيغة الرياضية المستخدمة في ذلك هي :

$$ر = \frac{مج(سَ - سَ)(ص - صَ)}{ن\ ع_س\ ع_ص}$$

حيث :

ر : ترمز إلى معامل الارتباط البسيط .

(س - سَ) : انحرافات قيم المتغير س عن وسطها الحسابي س

(ص - صَ) : انحرافات قيم المتغير ص عن وسطها الحسابي ص

ن : عدد المفردات للمتغير س ، ص .

$ع_س$: الانحراف المعياري لقيم المتغير س أي = $\sqrt{\frac{مج(س - سَ)^2}{ن}}$

$ع_ص$: الانحراف المعياري لقيم المتغير ص أي = $\sqrt{\frac{مج(ص - صَ)^2}{ن}}$

وهذا المعامل وضعه كارل بيرسون Carl Pearson ، وتنحصر قيمته بين +١ ، -١ ويكون معدوماً اذا ساوى الصفر .

ويمكن استنباط صيغة أسهل لمعامل الارتباط من الصيغة السابقة كما يلي:

$$\therefore مج(س-سَ)(ص-صَ) =$$

$$= مج(س\ ص - سَ\ ص - س\ صَ + سَ\ صَ)$$

---

* يرجع استخدام الانحرافات في عملية القياس الى أن أفضل طريقة لقياس التغير في العلاقة بين المتغيرين هو إيجاد الفرق بين قيم كل متغير ووسطه الحسابي .

** ويرجع استخدام القيم المعيارية في عملية القياس إلى التخلص من وحدات القياس المستخدمة في المتغيرين والتي قد لا تكون واحدة .



$$= \text{مجـ س ص} - \bar{\text{س}}\ \text{مجـ ص} - \bar{\text{ص}}\ \text{مجـ س} + \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}}$$

$$= \text{مجـ س ص} - \text{ن}\ \bar{\text{س}} \cdot \frac{\text{مجـ ص}}{\text{ن}} - \text{ن}\ \bar{\text{ص}} \cdot \frac{\text{مجـ س}}{\text{ن}} + \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}}$$

$$= \text{مجـ س ص} - \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}} - \text{ن}\ \bar{\text{ص}}\ \bar{\text{س}} + \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}}$$

$$= \text{مجـ س ص} - \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}}$$

$$= \text{مجـ س ص} - \text{ن} \times \frac{\text{مجـ س}}{\text{ن}} \times \frac{\text{مجـ ص}}{\text{ن}}$$

$$= \text{مجـ س ص} - \frac{\text{مجـ س} \cdot \text{مجـ ص}}{\text{ن}}$$

$$\therefore \text{ر} = \frac{\text{مجـ س ص} - \frac{\text{مجـ س} \cdot \text{مجـ ص}}{\text{ن}}}{\text{ن}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}}}$$

$$= \frac{\text{مجـ س ص} - \frac{\text{مجـ س} \cdot \text{مجـ ص}}{\text{ن}}}{\text{ن} \sqrt{\frac{\text{مجـ}(\text{س} - \bar{\text{س}})^2}{\text{ن}}} \sqrt{\frac{\text{مجـ}(\text{ص} - \bar{\text{ص}})^2}{\text{ن}}}}$$

$$= \frac{\text{مجـ س ص} - \frac{\text{مجـ س} \cdot \text{مجـ ص}}{\text{ن}}}{\sqrt{\text{مجـ}(\text{س} - \bar{\text{س}})^2} \sqrt{\text{مجـ}(\text{ص} - \bar{\text{ص}})^2}}$$

$$\therefore \text{ر} = \left(\frac{\text{مجـ س ص} - \frac{\text{مجـ س} \cdot \text{مجـ ص}}{\text{ن}}}{\sqrt{\left[\text{مجـ س}^2 - \frac{(\text{مجـ س})^2}{\text{ن}}\right]\left[\text{مجـ ص}^2 - \frac{(\text{مجـ ص})^2}{\text{ن}}\right]}} \right)$$

وهذه هي الصيغة النهائية الأسهل استخداما .



مثال (١)

من بيانات المثال رقم (١) فى الارتباط البسيط بيانيا أوجد معامل الارتباط البسيط ذو الخط المستقيم جبريا .

الحـــل

تكوين الجدول الاحصائى اللازم :

س	ص	س ص	$س^2$	$ص^2$
١١	٦	٦٦	١٢١	٣٦
٨	١٠	٨٠	٦٤	١٠٠
١٠	١٦	١٦٠	١٠٠	٢٥٦
١٧	١٤	٢٣٨	٢٨٩	١٩٦
١٤	٤	٥٦	١٩٦	١٦
٦٠	٥٠	٦٠٠	٧٧٠	٦٠٤

$$\therefore ر = \frac{\text{مج س ص} - \frac{\text{مج س . مج ص}}{ن}}{\sqrt{\left[\text{مج س}^{2} - \frac{\text{مج (س)}^{2}}{ن}\right]\left[\text{مج ص}^{2} - \frac{\text{مج (ص)}^{2}}{ن}\right]}}$$

$$= \frac{٦٠٠ - \frac{٦٠ \times ٥٠}{٥}}{\sqrt{\left[٧٧٠ - \frac{\times ٦٠}{٦٠}\right]\left[٦٠٤ - \frac{\times ٥٠}{٥٠}\right]}}$$

$$= \frac{٦٠٠ - ٦٠٠}{\sqrt{\left[٧٧٠ - \frac{\times ٦٠}{٦٠}\right]\left[٦٠٤ - \frac{\times ٥٠}{٥٠}\right]}}$$



$$= \frac{\text{صفر}}{\sqrt{\left[٧٧٠ - \frac{٦٠ \times ٦٠}{٦٠} \right] \left[٦٠٤ - \frac{٥٠ \times ٥٠}{٥٠} \right]}}$$

= صفـــر

التفسير:

بما أن معامل الارتباط الناتج (ر) يساوى الصفر ، فالارتبـــاط معدوم بين المتغيرين محل الدراسة ، وهذا ما سبق أن اوضحه العـــرض البيانى .

مثال (٢)

من بيانات المثال رقم (٢) فى الارتباط البسيط بيانيا أوجد معامل الارتباط البسيط ذو الخط المستقيم جبريا .



الحـل

تكوين الجدول الاحصائى اللازم :

س	ص	س ص	$س^2$	$ص^2$
١	٧	٧	١	٤٩
٣	١٩	٥٧	٩	٢٥٦
٤	٢٥	١٠٠	١٦	٦٢٥
٧	٤٣	٣٠١	٤٩	١٨٤٩
١٠	٦١	٦١٠	١٠٠	٣٧٢١
٢٥	١٥٥	١٠٧٥	١٧٥	٦٦٠٥

$$∴ ر = \frac{مجـ س ص - \frac{مجـ س . مجـ ص}{ن}}{\sqrt{\left[مجـ س^2 - \frac{مجـ (س)^2}{ن}\right]\left[مجـ ص^2 - \frac{مجـ (ص)^2}{ن}\right]}}$$

$$= \frac{١٠٦٦ - \frac{٢٥ \times ١٥٥}{٥}}{\sqrt{\left[١٧٥ - \frac{٢٥ \times ٢٥}{٥}\right]\left[٦٦٠٥ - \frac{١٥٥ \times ١٥٥}{٥}\right]}}$$

$$= \frac{١٠٦٦ - ٧٧٥}{\sqrt{١٧٥ - ١٨٠٠}}$$

$$= \frac{٣٠٦}{٣٠٦}$$

$$= ١$$

التفسير :

بما أن معامل الارتباط الناتج (ر) يساوى الواحد الصحيح ، فالارتباط تام بين المتغيرين محل الدراسة ، وبما أن الاشارة موجبة لهذا المعامل ، فالارتباط طردى وهذا ما سبق أن أوضحه العرض البيانى .

مثال (٣)

من بيانات المثال رقم (٣) فى الارتباط البسيط بيانيا أوجد معامل الارتباط البسيط ذو الخط المستقيم جبريا .

الحـــل

تكوين الجدول الاحصائى اللازم :

س	ص	س ص	$س^2$	$ص^2$
٢	١١	٢٢	٤	١٢١
٤	١٠	٤٠	١٦	١٠٠
٨	٨	٦٤	٦٤	٦٤
١٠	٧	٧٠	١٠٠	٤٩
١٦	٤	٦٤	٢٥٦	١٦
٤٠	٤٠	٢٦٠	٤٤٠	٣٥٠

$$\therefore ر = \frac{مجـ\, س\, ص - \frac{مجـ\, س \,.\, مجـ\, ص}{ن}}{\sqrt{\left[مجـ\, س^2 - \frac{(مجـ\, س)^2}{ن}\right]\left[مجـ\, ص^2 - \frac{(مجـ\, ص)^2}{ن}\right]}}$$



$$= \frac{٢٦٠ - \frac{٤٠ \times ٤٠}{٥}}{\sqrt{\left(٤٤٠ - \frac{٤٠ \times ٤٠}{٥}\right) \left(٣٥٠ - \frac{٤٠ \times ٤٠}{٥}\right)}}$$

$$= \frac{-٦٠}{\sqrt{(١٢٠)\ (٣٠)}}$$

$$= \frac{-٦٠}{٦٠}$$

$$= -١$$

التفسير:

بما أن معامل الارتباط الناتج يساوى الواحد الصحيح بإشارة سالبة ، فالارتباط تام عكسى ، وهذا ما أوضحه العرض البيانى .

مثال (٤)

من بيانات المثال رقم (٤) فى الارتباط البسيط بيانيا أوجد معامل الارتباط ذو الخط المستقيم جبريا .



الحـــل

تكوين الجدول الاحصائى اللازم :

س	ص	س ص	س^2	ص^2
٧٨	٥٥	٤٢٩٠	٦٠٨٤	٣٠٢٥
٦٥	٦١	٣٩٦٥	٤٢٢٥	٣٧٢١
٦٣	٥٠	٣١٥٠	٣٩٦٩	٢٥٠٠
٥٥	٥٣	٢٩١٥	٣٠٢٥	٢٨٠٩
٥٣	٢٨	١٤٨٤	٢٨٠٩	٧٨٤
٢٤	٣١	٧٤٤	٥٧٦	٩٦١
٤٣	٣٥	١٥٠٥	١٨٤٩	١٢٢٥
٤٦	٣٥	١٦١٠	٢١١٦	١٢٢٥
٤٨	٤٧	٢٢٥٦	٢٣٠٤	٢٢٠٩
٥٥	٤٥	٢٤٧٥	٣٠٢٥	٢٠٢٥
٥٣٠	٤٤٠	٢٤٣٩٤	٢٩٩٨٢	٢٠٤٨٤

$$\therefore \text{ر} = \frac{\text{مجـ س ص} - \dfrac{\text{مجـ س . مجـ ص}}{\text{ن}}}{\sqrt{\left[\text{مجـ س}^2 - \dfrac{(\text{مجـ س})^2}{\text{ن}}\right]\left[\text{مجـ ص}^2 - \dfrac{(\text{مجـ ص})^2}{\text{ن}}\right]}}$$

$$= \frac{٢٤٣٩٤ - \dfrac{٥٣٠ \times ٤٤٠}{١٠}}{\sqrt{\left[٢٩٩٨٢ - \dfrac{٥٣٠ \times ٥٣٠}{١٠}\right]\left[٢٠٤٨٤ - \dfrac{٤٤٠ \times ٤٤٠}{١٠}\right]}}$$

$$= \frac{١٠٧٤}{\sqrt{(١٨٩٢)\ (١١٢٤)}}$$



$$= \frac{١٠٧٤}{١٤٥٨,٣}$$

$$= ٠,٧٤$$

التفسير:

بما أن معامل الارتباط الناتج يساوى ٠,٧٤ بإشارة موجبة ، فالارتباط طردى قوى ، وهذا ما أوضحه العرض البيانى للعلاقة بين هذين المتغيرين ، وهنا يمكن القول أنه كلما قرب معامل الارتباط من الواحد الصحيح كلما كان انتشار النقط على جانبى خط الانتشار ، وأنه كلما بعد معامل الارتباط عن الواحد الصحيح كلما كان انتشار النقط بعيدا عن خط الانتشار .

مثال (٥)

من بيانات المثال رقم (٥) فى الارتباط البسيط بيانيا أوجد معامل الارتباط البسيط ذو الخط المستقيم جبريا .

الحل

تكوين الجدول الاحصائى اللازم :

س	ص	س ص	$س^٢$	$ص^٢$
١٢	١	١٢	١٤٤	١
١٠	٢	٢٠	١٠٠	٤
٩	٦	٥٤	٨١	٣٦
٦	٩	٥٤	٣٦	٨١
٣	١٨	٥٤	٩	٣٢٤
٤٠	٣٦	١٩٤	٣٧٠	٤٤٦

$$ر = \frac{\text{مج س ص} - \frac{\text{مج س . مج ص}}{\text{ن}}}{\sqrt{\left[\text{مج س}^{2} - \frac{(\text{مج س})^{2}}{\text{ن}}\right]\left[\text{مج ص}^{2} - \frac{(\text{مج ص})^{2}}{\text{ن}}\right]}}$$

$$= \frac{١٤٩ - \frac{٤٠ \times ٣٦}{٥}}{\sqrt{\left[٣٧٠ - \frac{٤٠ \times ٤٠}{٥}\right]\left[٤٤٦ - \frac{٣٦ \times ٣٦}{٥}\right]}}$$

$$= \frac{-٩٤}{\sqrt{٥٠ \times ١٨٦٫٨}}$$

$$= \frac{-٩٤}{٩٦٫٦٤٣}$$

$$= -٠٫٩٧$$

التفسير:

بما أن معامل الارتباط الناتج يساوى -٠,٩٧ أى بإشارة سالبة فالارتباط عكسى ، وحيث أن قيمته تقترب من الواحد الصحيح فالارتباط قوى ، وهذا ما أوضحه العرض البيانى للعلاقة بين هذين المتغيرين .

مغزى الارتباط التام وغير التام :

يحدث الارتباط التام اذا كان التغير فى أحد المتغيرين يتبعه تغير فى المتغير الآخر ويتم حسابه بالضبط كالعلاقة بين طول ضلع المربع ومساحته ، أو طول ضلع المكعب وحجمه .... وكما هو موجود فى المثالين رقمى ٢ ، ٣ وهذا النوع من الارتباط نادر الحدوث .

أما الارتباط غير التام فيحدث اذا كان التغير فى أحد المتغـــيرين يتبعه تغير فى المتغير الآخر ولا يمكن حسابه بالضبط وإنما لحـــد مـــا ، وهذا النوع من الارتباط هو الشائع بين الظواهر .

مثال لحالـــــه

أوضحنا فيما سبق كيفية حساب معامل الارتباط لعدد قليـــل مـــن القيم (مجتمع احصائى صغير) ، إلا أنه فى حالة العدد الكبير من القيـــم (مجتمع احصائى كبير) فإن حساب معامل الارتباط يصبح اكثر تعقيـدا ، ولذلك فالامر يستلزم وضع هذه البيانات فى جدول توزيع تكرارى ليسهل حساب معامل الارتباط كما فى المثال التالى :

الجدول التالى يبين الحالة الدراسية لمجموعة من ٦٩ طـــالب بالســياحة والفنادق فى مادتى الاقتصاد والاحصاء (متغيرين) :

الاقتصاد	الاحصاء	الاقتصاد	الاحصاء	الاقتصاد	الاحصاء	الاقتصاد	الاحصاء
١٥	٥	٢٨	١٧	٣٥	١٥	٤٢	٢٩
٢٠	٨	٢٨	١٨	٣٦	١٦	٤٣	٣٢
٢٢	٨	٢٩	١٩	٣٧	١٧	٤٤	٣٣
٢٤	١٨	٣٠	٢٠	٣٨	١٨	٤٥	١٥
٢٥	٥	٣١	٢١	٣٩	٢٦	٤٦	١٦
٢٦	١٥	٣٢	٢٤	٤٠	٢٧	٤٧	١٧
٢٧	١٦	٣٤	٣٠	٤١	٢٨	٤٨	١٨
٤٩	٢٠	٥٦	٢٥	٦٥	١٦	٧٣	٣٥
٥٠	٢٥	٥٧	٢٧	٦٦	١٧	٧٤	٣٦
٥١	٢٦	٥٨	٢٨	٦٧	٢٥	٧٤	٣٧
٥٢	٢٧	٥٩	٣٠	٦٨	٢٦	٦٧	٣٨

تابع الجدول

الاقتصاد	الاحصاء	الاقتصاد	الاحصاء	الاقتصاد	الاحصاء	الاقتصاد	الاحصاء
٥٣	٢٨	٦٠	٣٥	٦٩	٢٧	٦٧	٣٩
٥٤	٢٩	٦١	٣٦	٧٠	٢٨	٧٠	٤٠
٥٤	٣٠	٦٢	٣٧	٧٠	٢٩	٧٠	٤٥
٥٣	٣١	٦٣	٣٨	٧١	٣٠	٧١	٥٤
٥٢	٣٥	٦٤	٤٠	٧٢	٣١		
٥٠	٣٨	٦٤	٥٠	٧٢	٣٢		
٥٥	٢٠	٦٥	١٥	٧٣	٣٥		

المطلوب :

حساب معامل الارتباط بين متغيرى الاقتصاد والاحصاء لمجموعة الطلاب محل الدراسة .

الحل

يتطلب الحل تكوين الجداول الاحصائية اللازمة التالية :

تكوين الجدول الاحصائى اللازم أ (التوزيع التكرارى المزدوج) :

الاحصاء / الاقتصاد	(٥-١٥)	(١٥-٢٥)	(٢٥-٣٥)	(٣٥-٤٥)	(٤٥-٥٥)	مجموع تكرارات الاقتصاد
(١٥ ـ ٢٥)	٣	١				٤
(٢٥ ـ ٣٥)	١	٨	١			١٠
(٣٥ ـ ٤٥)		٤	٦			١٠
(٤٥ ـ ٥٥)		٥	٧	٢		١٤
(٥٥ ـ ٦٥)		١	٤	٥	١	١١
(٦٥ ـ ٧٥)		٣	٨	٧	٢	٢٠
مجموع تكرارات الاحصاء	٤	٢٢	٢٦	١٤	٣	٦٩



تكوين الجدول الاحصائى اللازم ب (التوزيع الهامشى لمتغير الاقتصاد):

فئات متغير الاقتصاد	التكرار (ك)	مركز الفئة (س)	ك س	ك س^{٢}
(١٥ – ٢٥)	٤	٢٠	٨٠	١٦٠٠
(٢٥ – ٣٥)	١٠	٣٠	٣٠٠	٩٠٠٠
(٣٥ – ٤٥)	١٠	٤٠	٤٠٠	١٦٠٠٠
(٤٥ – ٥٥)	١٤	٥٠	٧٠٠	٣٥٠٠٠
(٥٥ – ٦٥)	١١	٦٠	٦٦٠	٣٩٦٠٠
(٦٥ – ٧٥)	٢٠	٧٠	١٤٠٠	٩٨٠٠٠
المجموع	٦٩		٣٥٤٠	١٩٩٢٠٠

تكوين الجدول الاحصائى اللازم جـ(التوزيع الهامشى لمتغير الاحصاء)

فئات متغير الاحصاء	التكرار (ك)	مركز الفئة (ص)	ك ص	ك ص^{٢}
(٥ – ١٥)	٤	١٠	٤٠	٤٠٠
(١٥ – ٢٥)	٢٢	٢٠	٤٤٠	٨٨٠٠
(٢٥ – ٣٥)	٢٦	٣٠	٧٨٠	٢٣٤٠٠
(٣٥ – ٤٥)	١٤	٤٠	٥٦٠	٢٢٤٠٠
(٤٥ – ٥٥)	٣	٥٠	١٥٠	٧٥٠٠
المجموع	٦٩		١٩٧٠	٦٢٥٠٠

تكوين الجدول الاحصائى اللازم له (التوزيع التكرارى المزدوج لتكوين الخلايا):

ص / س	١٠	٢٠	٣٠	٤٠	٥٠	مجموع الخلايا افقيا
٢٠	٣ ٦٠٠	١ ٤٠٠				١٠٠
٣٠	١ ٣٠٠	٨ ٤٨٠٠	١ ٩٠٠			٦٠٠٠
٤٠		٤ ٣٢٠٠	٦ ٧٢٠٠			١٠٤٠٠
٥٠		٥ ٥٠٠٠	٧ ١٠٥٠٠	٢ ٤٠٠٠		١٩٥٠٠
٦٠		١ ١٢٠٠	٤ ٧٢٠٠	٥ ١٢٠٠	١ ٣٠٠٠	٢٣٤٠٠
٧٠		٣ ٤٢٠٠	٨ ١٦٩٠٠	٧ ١٩٦٠٠	٢ ٧٠٠٠	٤٧٦٠٠
مجموع الخلايا رأسيا	٩٠٠	١٨٨٠٠	٤٢٦٠٠	٣٥٦٠٠	١٠٠٠٠	١٠٧٩٠٠

وقيمة الخلية = س × ص × (ك المناظرة عند تقاطع الصف والعمود) ،

فمثلا :

الخلية ٦٠٠ = س × ص × (ك المناظرة عند تقاطع الصف والعمود)

= ٢٠ × ١٠ × ٣ = ٦٠٠

الخلية ١٠٥٠٠ = س + ص + (ك المناظرة عند تقاطع الصف والعمود)

= ٥٠ × ٣٠ × ٧ = ١٠٥٠٠ وهكذا

ويتضح من الجدول أن الرقم ١٠٧٩٠٠ هو مجموع الخلايا سواء رأسيا أو افقيا وهو عبارة عن (مج س ص ك) .

وحيث أن معامل الارتباط فى هذه الحالة يساوى :



$$ر.. = \frac{مج\ س\ ص\ ك - \frac{مج\ س\ ك\ .\ مج\ ص\ ك}{ن}}{\sqrt{\left[مج\ س^{٢}ك - \frac{مج(س\ ك)^{٢}}{ن}\right]\left[مج\ ص^{٢}ك - \frac{مج(ص\ ك)^{٢}}{ن}\right]}}$$

$$= \frac{١٠٧٩٠٠ - \frac{٣٥٤٠ - ١٩٧٠}{٦٩}}{\sqrt{\left[١٩٩٢٠٠ - \frac{(٣٥٤٠)^{٢}}{٦٩}\right]\left[٦٢٥٠٠ - \frac{(١٩٧٠)^{٢}}{٦٩}\right]}}$$

$$= \frac{٦٨٣٠,٤٣}{\sqrt{١٧٥٨٢,٦١ \times ٦٢٥٥,٠٧}}$$

$$= ٠,٧$$

التفسير:

بما أن معامل الارتباط موجب ، فالارتباط طـــردى ، وبمـــا أن معامل الارتباط ٠,٧ فالارتباط قوى . أى أن مادتى الاقتصاد والاحصـاء بينهما علاقة طردية قوية لطلبة السياحة والفنادق .

مثال لحالــــة

قد يكون المتغيرين محل الدراسة بياناتها غير عددية أى قياســها غير كمى أى قياسا وصفيا ، ونريد دراسة العلاقة بينها كما فى المثــال التالى :



مجموعة مكونة من ١٠٠ فرد يعملون فى مجال الفنادق وقد تم جمع بيانات من كل منهم عن نوعه (ولد أم بنت) وعن نوع عمله (فى المطبخ أم فى المكاتب الامامية) ، وقد تم عرض البيانات التى تم جمعها من كل منهم فى جدول التوزيع التكرارى المزدوج التالى :-

نوع العمل / نوع الفرد	فى المطبخ	فى المكاتب الامامية	المجموع
ولد	٣٥	٢٥	٦٠
بنت	٣٠	١٠	٤٠
المجموع	٦٥	٣٥	١٠٠

والمطلوب :

هل يوجد علاقة بين متغيرى نوع الفرد ونوع العمل لمجموعة هؤلاء الأفراد .

الحل

يلاحظ على جدول التوزيع التكرارى المزدوج فى هذه الحالة أن كلا من المتغيرين مقسم الى فئتين وصفيتين ، وهنا لدراسة العلاقة بين هذين المتغيرين يتم استخدام معامل يسمى "معامل الاقتران" assosiation وليس بمعامل الارتباط السابق شرحه .

تكوين الجدول الاحصائى اللازم (جدول الاقتران) :

نوع العمل / نوع الفرد	فى المطبخ		فى المكاتب الامامية	
ولـــد	٣٥	أ	٢٥	ب
بنت	٣٠	جـ	١٠	د

.. معامل الاقتران (ق) = $\frac{\text{أ د} - \text{ب جـ}}{\text{أ د} + \text{ب جـ}}$

.. ق = $\frac{٣٥ \times ١٠ - ٢٥ \times ٣٠}{٣٥ \times ١٠ + ٢٥ \times ٣٠}$ = ٠,٣٦

وهذا معناه أن هناك علاقة بين نوع الفرد (ولد أم بنت) ونـــوع العمل فى الفنادق (فى المطبخ أم فى المكاتب الامامية) وإلا كان معـامل الاقتران (ق) يساوى الصفر ، وبديهى أن معامل الاقتران أقل من الواحـد الصحيح .

مثـــال لحالـــــة

قد يكون المتغيرين محل الدراسة بياناتها كما فى حالـــــة المثـال السابق إلا أن جدول التوزيع التكرارى المزدوج نجد فيه كلا المتغـــيرين مقسم الى اكثر من فئتين وصفيتين كما فى المثال التالى :

تقدم لامتحان المقابلة الشخصية للالتحاق بالسياحة والفنــادق ١٤٤ طالب من ثلاثة أنواع من المدارس (ثانوي عام ، ثانوي فندقى ، ثـــانوى تجارى) وكانت نتيجة الامتحان كما فـــى جـــدول التوزيـــع التكـرارى

المزدوج التالى :

درجة الالتحاق / نوع المدرسة	٠ - ٢٠	٢٠-٤٠	٤٠-٦٠	٦٠-٨٠	٨٠-١٠٠	المجموع
ثانوي عام	٣٤	٢٢	٨	٢	٢	٦٨
ثانوي فندقي	صفر	١	٩	١٨	١	٢٩
ثانوي تجارى	١٦	٢٧	٢	صفر	٢	٤٧
المجموع	٥٠	٥٠	١٩	٢٠	٥	١٤٤

والمطلوب :

هل يوجد علاقة بين نوع المدرسة ونتيجة الالتحاق بالسياحة والفنادق

الحـل :

لدراسة العلاقة بين هذين المتغيرين يتم استخدام معامل آخر يسمى معامل التوافق Contingeny .

تكوين الجدول الاحصائى اللازم (جدول التوافق) :

درجة الامتحان / نوع المدرسة	٠-٢٠	٢٠-٤٠	٤٠-٦٠	٦٠-٨٠	٨٠-١٠٠	المجموع
ثانوي عام	١١٥٦ / ٦٨×٥٠	٤٨٤ / ٦٨×٥٠	٦٤ / ٦٨×١٩	٤ / ٦٨×٢٠	٤ / ٦٨×٥	٠,٥٦
ثانوي فندقي	صفر / ٢٩×٥٠	١ / ٢٩×٥٠	٨١ / ٢٩×١٩	٣٢٤ / ٢٩×٢٠	١ / ٢٩×٥	٠,٧١
ثانوي تجاري	٢٥٦ / ٤٧×٥٠	٧٢٩ / ٤٧×٥٠	٤ / ٤٧×١٩	صفر / ٤٧×٢٠	٤ / ٤٧×٥	٠,٤٤
المجموع	٠,٤٥	٠,٤٥	٠,٢٠	٠,٥٦	٠,٠٤	١,٧

وتتكون قيمة كل خلية فى هذا الجدول من مربع تكـــرار تقـــاطع الصف والعمود فى الجدول الاصلى مقسوما على حاصل ضرب مجمـوع تكرار هذا الصف فى مجموع تكرار هذا العمود ، فمثلا خليـــة الثـــانوي العام عند الفئة (٠ — ٢٠) أى $\frac{١١٥٦}{٦٨ \times ٥٠}$ هى حاصل :

$$\frac{٣٤ \times ٣٤}{\text{مجموع تكرار الصف (٦٨)} \times \text{مجموع تكرار العمو (٥٠)}}$$ وهكذا .

، .. معامل التوافق = $\sqrt{\frac{\text{جـ} - ١}{\text{جـ}}}$

حيث :

جـ : ترمز الى المجموع الكلى فى جدول التوافق

.. معامل التوافق = $\sqrt{\frac{١,٧ - ١}{١,٧}}$ = ٠,٦٤

وهذا معناه أن هناك علاقة بين نوع المدرسة ومــــدى الالتحـــاق بالسياحة والفنادق .

مثـــــال

قد يكون المتغيرين محل الدراسة بياناتها وصفية لكن عن مجتمـــع احصائى صغير كما فى المثال التالى :



مجموعة مكونة من عشرة طلاب ودرجات نجاحها في مادتي الاحصاء والاقتصاد كما في الجدول التالي :

التقدير في الاحصاء	جيد	ممتاز	جيد	مقبول	ضعيف	جيد	ضعيف جدا	ضعيف	جيد جدا	جيد جدا
التقدير في الاقتصاد	جيد	مقبول	جيد جدا	جيد	ضعيف جدا	جيد جدا	ضعيف	مقبول	ممتاز	جيد جدا

والمطلوب :

هل يوجد علاقة بين هذين المتغيرين ؟

الحـل

لقياس العلاقة بين هذين المتغيرين يتم استخدام مقياس آخر للارتباط يسمى بمعامل ارتباط الرتب (سبيرمان) ، وفيه يتم ترتيب كل من المتغيرين حسب القيم بمعنى أنه تعطى أكبر القيم وصفا الرتبة ١ ، والقيمة الأقل مباشرة الرتبة ٢ ، .... وهكذا ، وفي حالة القيم المتساوية فإنها تأخذ نفس الرتبة على اساس المتوسط لمجموع رتب هذه القيم المتساوية ، ثم يتم التعويض في القانون التالي (سبيرمان) :

$$ر = ١ - \frac{٦ \text{ مج } ف^{٢}}{ن \left(ن^{٢} - ١ \right)} \quad (*)$$

(*) يمكن من معامل ارتباط بيرسون أن نشتق معامل الارتباط (سبيرمان)



وعليـــه

تكوين الجدول الاحصائى اللازم :

التقدير فى الاحصاء س	التقدير فى الاحصاء ص	رتب س	رتب ص	فروق الرتب ف	مربع الفروق $ف^2$
جيد	جيد	٥	٥.٥	٠,٥	٠,٢٥
ممتاز	مقبول	١	٧,٥	٦,٥	٤٢,٢٥
جيد	جيد جدا	٥	٣	٢,٥	٤,٠٠
مقبول	جيد	٧	٥.٥	١,٥	٢,٢٥
ضعيف	ضعيف جدا	٨,٥	١٠	١,٥	٢,٢٥
جيد	جيد جدا	٥	٣	٢,٠٠	٤,٠٠
ضعيف جدا	ضعيف	١٠	٩	١.٠٠	١,٠٠
ضعيف	مقبول	٨,٥	٧,٥	١,٠٠	١,٠٠
جيد جدا	ممتاز	٢,٥	١	١,٥	٢,٢٥
جيد جدا	جيد جدا	٢,٥	٣	٠,٥	٠,٢٥
					مجموع = ٥٩,٥

ويلاحظ على الجدول السابق ما يلى :

- أن ترتيب قيم المتغير س (التقدير فى مادة الاحصاء) قد أخذ التقدير ممتاز الرتبة رقم ١ .

- أن الرتبة الثانية والثالثة قيمتها متساوية أى جيد جدا وفى هذه الحالة يتم أعطاء هاتين المفردتين نفس الرتبة على اساس متوسط مجموع رتبتيهما أى $\frac{٢ + ٣}{٢}$ = ٢,٥ وبالتالى تصبح الرتبة الثانية والثالثة لا وجود لهما .

- أن الرتبة الرابعة والخامسة والسادسة قيمهم متساوية أى التقدير جيد ، وفى هذه الحالة يتم أعطاء هذه المفردات نفس الرتبة علـــى اســاس



متوسط مجموع ترتيبهم أى $\frac{٤ + ٥ + ٦}{٣}$ = ٥ وبالتالى تصبح الرتبة الرابعة والخامسة والسادسة لا وجود لها ، وهكذا .

- ويمكن اجراء عملية الترتيب السابقة كما يلى :

الترتيب	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
توزيع التقدير	م	جج	جـ	جـ	جـ	جـ	ل	ض	ض	ض جـ
توزيع الرتب	١	٢,٥	٢,٥	٥	٥	٥	٧	٨,٥	٨.٨	١٠

- ويتم ترتيب قيم المتغير ص (التقدير فى مادة الاقتصاد) كما هو الحال فى ترتيب قيم المتغير س .

- بعد أجراء عملية الترتيب للمتغيرين يتم استخدام مجموع الفرق بين كل رتبتين متناظرتين لقياس الارتباط بينهما ، وذلك لأن هذه الفووق تتوقف قيمتها على مدى الاتفاق أو الاختلاف بين الرتب المناظرة ، ونظرا لكون مجموع هذه الفروق قد يساوى الصفر (حيث إشارات بعضها موجب والبعض الأخر سالبا) رغم وجود فروق ، فأنه يتم تربيع هذه الفروق ثم التعويض بمجموع هذا التربيع فى القانون المراعى لهذه المعالجات .

$$ر = ١ - \frac{٦ \text{ مج } ف^٢}{ن (ن^٢ - ١)}$$

$$= ١ - \frac{٦ \times ٥٩,٥}{١٠ (٩٩)}$$

$$= ١ - ٠,٣٤ = ٠,٦٦$$

وهذا يعنى أن الارتباط قوى موجب (طردى) مما يدل على وجود علاقة بين المتغيرين (التقدير فى الاحصاء ، والتقدير فى الاقتصاد) .



ملاحظات :-

١-اذا كان اكبر رتب المتغير س تناظرها اكبر رتب المتغير ص ، والرتبة الثانية فى الكبر من رتب س تناظرها الرتبة الثانية فى الكبر من رتب ص ... وهكذا حتى نصل الى أصغر رتبة من رتب س التى تناظرها أصغر رتبة من رتب ص ، فإن الارتباط بين هذين المتغيرين تام موجب حيث :

الترتيب	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
رتب س	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
رتب ص	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
الفرق	٠	٠	٠	٠	٠	٠	٠	٠	٠	٠	
مربع الفرق	٠	٠	٠	٠	٠	٠	٠	٠	٠	٠	المجموع = صفر

$$\therefore \rho = ١ - \frac{٦ \text{ مج ف}^{٢}}{\text{ن}(\text{ن}^{٢} - ١)}$$

$$\therefore \rho = ١ - \frac{٠}{١٠(٩٩)}$$

$$= ١$$

وهذا يعنى أن الارتباط تام موجب .

٢-اذا كان أكبر رتب المتغير س تناظرها أصغر رتب المتغير ص ، والرتبة الثانية فى الكبر من رتب س تناظرها الرتبة الثانية فى الصغر من رتب ص .... وهكذا حتى تصل إلى أصغر رتبة من رتب س التى تناظرها أكبر رتبة من رتب ص ، فإن الارتباط بين هذين المتغيرين تام سالبة حيث :



الترتيب	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
رتب س	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
رتب ص	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١	
الفرق	-٩	-٧	-٥	-٣	-١	١	٣	٥	٧	٩	المجموع = صفر
مربع الفرق	٨١	٤٩	٢٥	٩	١	١	٩	٢٥	٤٩	٨١	المجموع = ٣٣٠

$$\therefore \text{ر} = ١ - \frac{٦ \text{ مج ف}^{٢}}{\text{ن} (\text{ن}^{٢} - ١)}$$

$$= ١ - \frac{١٩٨٠}{٩٩٠}$$

$$= ١ - \frac{٣٣٠ \times ٦}{١٠(٩٩)}$$

$$= ١ - ٢$$

$$= - ١$$

وهذا يعنى أن الارتباط تام سالب .

٣-أن أغلب التوزيعات لقيم متغيرين معا تقع بين الارتباط التام الموجب والارتباط التام السالب ، أى نادرا ما نجد ارتباطا بين متغيرين يكون ارتباط تام موجب أو ارتباط تام سالب .

٤-بالرغم من أن معامل ارتباط الرتب (سبيرمان) يمتاز بالسهولة إلا أنه أقل كفاءة من معامل ارتباط بيرسون وذلك لأن الأول لا يتعامل مع البيانات الأصلية للمتغيرين وإنما يتعامل مع ترتيبها ، وبصفة عامة يعتبر معامل ارتباط الرتب مناسبا بشكل خاص عند معالجة البيانات



الوصفية كدراسة العلاقة بين تقديرات الطلاب فى مادتين ، أو دراسة العلاقة بين رائحة الزهور وألوانها ، أو العلاقة بين الأطعمة ومذاقها.

٥-قد تكون بيانات المتغيرين محل الدراسة مقاسة قياسا كميا ، ومن ثم يتم دراسة العلاقة بينهما باستخدام معامل ارتباط (بيرسون) ، إلا أنه يفضل استخدام معامل ارتباط الرتب (سيبرمان) وذلك بغرض تخفيف العمليات الحسابية .

٦-وخلاصة العلاقة بين المتغيرين :

- إذا كانت قيم المتغيرين مقاسة قياسا كميا (فى صورة رقمية)، فإنه يتم دراسة العلاقة بينهما باستخدام معامل الارتباط البسيط Carlition .
- اذا كانت قيم المتغيرين مقاسة قياسا كميا وفى جدول توزيع تكرارى مزدوج فإنه يتم دراسة العلاقة بينهما باستخدام معامل الارتباط المعدل بالتكرارات .
- اذا كانت قيم المتغيرين غير مقاسة قياسا كميا أى مقاسة قياسا وصفيا وفى جدول توزيع تكرارى مزدوج (٢×٢) ، فإنه يتم دراسة العلاقة بينهما باستخدام معامل الاقتران assosiation
- اذا كانت قيم المتغيرين غير مقاسة قياسا كميا أى مقاسة قياسا وصفيا وفى جدول توزيع تكرارى مزدوج (ن × ن) ، فإنه يتم دراسة العلاقة بينهما باستخدام معامل التوافق Contingencg .
- اذا كانت قيم المتغيرين مقاسة قياسا وصفيا أو كميا وفى غير جدول توزيع تكرارى مزدوج (مجتمع احصائى صغير) ، فإنه يتم دراسة العلاقة بينها باستخدام معامل ارتباط الرتب (سيبرمان) .



- اذا كان شكل الانتشار لا يظهر العلاقة بين المتغيرين محـــل الدراسة فى شكل خط مستقيم وانما يظهرهما فى شكل خـــط منحنى ، فإن قياس هذه العلاقة لا يتـــم باســـتخدام مقـــاييس الارتباط السابقة ، وإنما يتم باستخدام :

أ- دليل الارتباط Carrelation Index وذلك فـــى حالة البيانات غير المبوبـــة (المجتمـــع الاحصـــائى الصغير) .

ب- نسبة الارتباط Carrelation Ratio وذلك فـــى حالة البيانات المبوبة (المجتمع الاحصائى الكبير) .

وسيتم تأجيل دراسة هذين المقياسين الى مواضع آخرى .



الفصل الثانى

الانحدار البسيط

مقدمه :-

سبق ايضاح أن الباب الرابع يتعرض لدراسة العلاقة بين متغيرين فى مجموعة واحدة ، وقد تم فى الفصل السابق دراسة هذه العلاقة من خلال الارتباط ، ويتبين أن الارتباط فى قياسه لهذه العلاقة يبين درجة هذه العلاقة بمعنى هل هى قوية أم ضعيفة ، كما يبين فى نفس الوقت نوع هذه العلاقة بمعنى هل هى طردية (موجبة) أم عكسية (سالبة) .

والآن فى هذا الفصل نتعرض لدراسة هذه العلاقة أيضا لكن من خلال الانحدار ، وذلك الذى يبحث فى سببية هذه العلاقة بمعنى أن العلاقة لابد وأن يتحدد فيها المتغير التابع والمتغير المستقل ، ومن المعلوم أن المتغير المستقل هو الذى يؤثر فى المتغير التابع . ويسعى الانحدار من خلال الرياضيات والاحصاء واستخدام البيانات التى تم جمعها عن المتغيرين محل الدراسة التوصل الى المعادلة الرياضية التى تحكم سير هذه العلاقة ، ومن ثم نستطيع من خلال هذه المعادلة الرياضية معرفة درجة تأثير المتغير المستقل على المتغير التابع هذا بالاضافة إلى امكانية التنبؤ بما سيكون عليه المتغير التابع فى المستقبل إذا ما أخذ المتغير المستقل فى المعادلة الرياضية عامل الزمن .

وتجدر الاشارة إلى أن العلاقة الانحدارية قد تكون بين متغير تابع وتغير مستقل واحد وهى ما تعرف بالانحدار البسيط ، وقد تكون بين

متغير تابع وعدة متغيرات مستقلة وهو ما يعرف بالانحدار المتعدد وهذا الأخير سيتم تناوله فى الباب القادم . والانحدار البسيط قد يظهره شكل الانتشار على هيئة خط مستقيم أو على هيئة خط منحنى بدرجاته المختلفة .

الانحدار البسيط ذو الخط المستقيم :-

مثال

مجموعة مكونة من خمس طلاب بدرجات جاحهما فى مادتى الاقتصاد والاحصاء كما فى الجدول التالى :

درجات الأحصاء (س)	٢	٣	٥	٧	٨
درجات الاقتصاد (ص)	٢	٤	٦	٦	٧

والمطلوب :

١-ارسم شكل الانتشار .

٢-ايجاد خط الانتشار من على الرسم بمجرد النظر موضحا رأيك فى هذا الخط .

٣-ايجاد خط الانتشار رياضيا على اعتبار أنه خط انحدار ص/س ، ووضحه بيانيا .

٤- أوجد الفروق بين قيم ص الفعلية وهى الموجودة بالجدول الاصلى وقيم ص المقدرة وهى الناتجة من معادلة خط الانحدار المتحصل عليه ، ثم وضح هذه الفروق بيانيا ، وما رأيك فى فى مجموع هذه الفروق ، وأيضا مجموع مربعاتها.



٥-اختبر دقة التقدير لخط الانحدار الناتج ، لاجراء هذا الاختبار استخدم الآتى :-

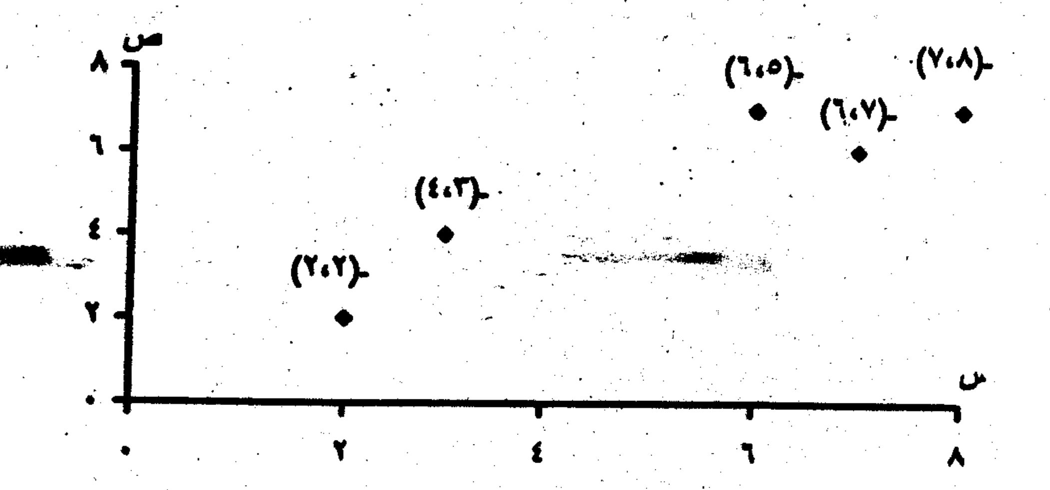
- الخطأ المعيارى لتقدير خط الانحدار ص/س .
- الخطأ المعيارى لتقدير معامل خط الانحدار ص/س .
- اختبار (ت) معنوية معامل الانحدار .

٦-بعد اجراء الاختبارات السابقة فما رأيك النهائى فى استخدام المعادلة الرياضية المتحصل عليها فى عملية التقدير .

الاجابة

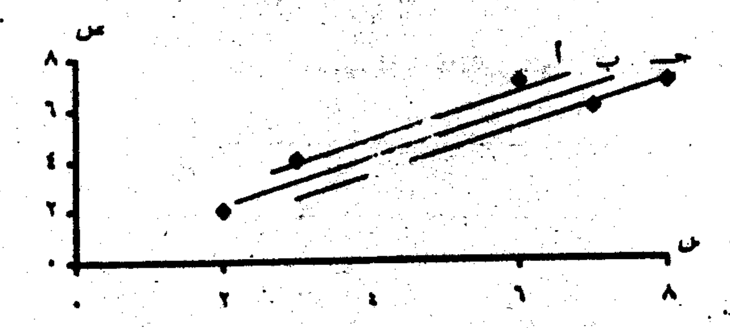
١- شكل الانتشار :

ويتكون شكل الانتشار من ٥ نقط الموضحة على الرسم وهى (٢ ، ٢) ، (٣ ، ٤) ، (٥،٦) ، (٧،٦) ، (٨،٧) .





٢- تقدير الخط المستقيم لانحدار ص/س وبمجرد النظر :



طالما خط الانتشار يمكن ان يتحدد بمجرد النظر فإنه سيختلف من شخص لآخر ، فقد يحدده شخص على أنه الخط أ ، وقد يحدده شخص آخر على أنه الخط ب ، وآخر على أنه الخط جـ . وهكذا . ويتحدد خط الانتشار بمجرد النظر ى بالتمهيد باليد عن طريق رسم خط يمر بأكبر عدد من نقط الانتشار وعلى أن يمر بين النقط الأخرى بالتوازن مع اهمال النقاط الشاذة أن وجدت . وهذه الطريقة تقريبية وتعوذها الدقة بل وتختلف من شخص لآخر ، ولذلك لاتستخدم فى العمل الاحصائى .

٣- تقدير الخط المستقيم لانحدار ص/س رياضيا :

أولا : تمهيد رياضى :

من المعلوم أنه اذا كانت الدالة ص = د(س) تخضع لمعادلة من الدرجة الاولى ص = أ + ب س ، فإنه الخط البيانى المعبر عن هذه الدالة هو الخط المستقيم .

مثال للايضاح

البيانات التالية عن متغيرين س ، ص معا فى شكل نقط .

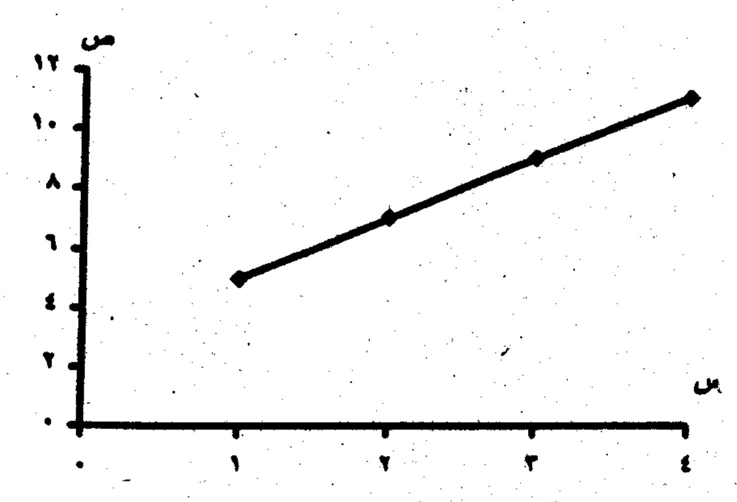
س	١	٢	٣	٤
ص	٥	٧	٩	١١

والمطلوب :

- عرض بيانات الجدول بيانيا واستنتاج الخط البيانى الناتج .
- ايجاد المعادلة الرياضية التى تحكم هذه العلاقة .

الحـــل

- العرض البيانى :



ومن الرسم يتضح أن الخط البيانى الناتج هو الخط المستقيم .

- إيجاد المعادلة الرياضية (*) التى تحكم هذه العلاقة :

(*) المعادلة الرياضية سواء خط مستقيم أو منحنى هى على وجه العموم علاقة جبريــة تعبر بالرموز عن المحل الهندسى لنقطه تتحرك بشروط معينة .



.. ب = $\frac{\Delta ص}{\Delta س}$ بين أى نقطتين من نقط الخط المستقيم

.. ب = $\frac{٧ - ٥}{٢ - ١}$ = ٢ أو = $\frac{٩ - ٧}{٣ - ٢}$ = ٢ أو $\frac{١١ - ٩}{٤ - ٣}$ = ٢

وهذا يتفق مع الحقيقة الهندسية إن ميل الخط المستقيم ثابت .

، .. أ = ص – ب س عند أى نقطة من نقط الخط المستقيم

.. أ = ٥ – ٢ × ١ = ٣ أو = ٧ – ٢ × ٢ = ٣ أو ...

.. ص = ٣ + ٢ س وهذه هى المعادلة الرياضية التى تحكم العلاقة بين المتغيرين س ، ص .

ثانيا : الرجوع إلى تقدير الخط المستقيم لانحدار ص/س رياضيا :

أفاد التمهيد الرياضى فى أنه اذا حصلنا على ب ، أ فإنه يمكن تحديد المعادلة الرياضية التى تحكم العلاقة المستقيمة بين المتغيرين س ، ص ، إلا أنه فى الاحصاء نظرا لكون نقاط الرسم البيانى ليست بالضرورة هى نقاط خط مستقيم بل هى فى الغالب نقاط تنتشر بعدا وقربا للخط المستقيم ، فإننا نسعى الى التوصل الى أفضل خط مستقيم يحكم العلاقة بين المتغيرين ، والأفضل خط هذا هو الخط الذى يتوسط نقط شكل الانتشار اذ عنده يكون مجموع مربعات انحرافات (فروق) نقاط شكل الانتشار عنه أقل ما يمكن(*) ، أى مجـ $(\hat{ص} - ص)^2$ يكون أقل ما يمكن حيث :

(*) تسمى بطريقة المربعات الصغرى Lest square Melhoud ، وتوجد طريقة أخرى لتقدير هذا الخط تسمى بمعادلة نيوتن للفروق المجزئة Newten's divided differences Formula



ضَ : هى ص التقديرية وهى = أ + ب س المقدر

، ص : هى ص الفعلية أى الأصلية والموجودة فى الجدول الأصلى للبيانات .

، مج (ضَ-ص)^{٢} : هو مقدار مجموع مربعات الفروق .

وإذا فرضنا أن مج (ضَ-ص)^{٢} تساوى ف

.. ف = مج (ضَ-ص)^{٢}

= مج (أ + ب س – ص)^{٢}

ولكى نحصل على أ ، ب اللتان تجعلان المقدار ف أقل ما يمكن ومن ثم الحصول على أفضل خط مستقيم لانحدار ص/س ، فإنه سيتم اجراء التفاضل الأول للدالة ف بالنسبة لـ أ ومساواة نتيجة هذا التفاضل بالصفر فنحصل على معادلة أ ، وهكذا بالنسبة لـ ب فنحصل على معادلة ب .

.. ف = مج (أ + ب س – ص)^{٢}

.. $\frac{د ف}{د أ}$ = ٢ مج (أ + ب س – ص)

= ٢ ن أ + ٢ ب مج س – ٢ مج ص

.. ٢ ن أ + ٢ب مج س – ٢ مج ص = صفر

(وبقسمة طرفى المعادلة على ٢)

.. مج ص = ن أ + ب مج س (١) (وبقسمة طرفى المعادلة على ن)

$$\therefore \text{أ} = \frac{\text{مجـ ص} - \text{ب مجـ س}}{\text{ن}} = \bar{\text{ص}} - \text{ب}\ \bar{\text{س}}$$

$$، \frac{\text{د ف}}{\text{د ب}} = ٢ \text{ مجـ} (\text{أ} + \text{ب س} - \text{ص}) \times \text{س}$$

$$= ٢ \text{ أ مجـ س} + ٢ \text{ ب مجـ س}^{٢} - ٢ \text{ مجـ س ص}$$

$$\therefore ٢ \text{ أ مجـ س} + ٢ \text{ ب مجـ س}^{٢} - ٢ \text{ مجـ س ص} = \text{صفر}$$

$$\therefore \text{أ مجـ س} + \text{ب مجـ س}^{٢} - \text{مجـ س ص} = \text{صفر}$$

$$\therefore \text{مجـ س ص} = \text{أ مجـ س} + \text{ب مجـ س}^{٢} \qquad (٢)$$ (*)

$$= \frac{\text{مجـ ص} - \text{ب مجـ س}}{\text{ن}} \times \text{مجـ س} + \text{ب مجـ س}^{٢}$$

$$= \frac{\text{مجـ س . مجـ ص}}{\text{ن}} - \text{ب} \frac{(\text{مجـ س})^{٢}}{\text{ن}} + \text{ب مجـ س}^{٢}$$

$$= \frac{\text{مجـ س . مجـ ص}}{\text{ن}} + \text{ب} \left[(\text{مجـ س})^{٢} - \frac{(\text{مجـ س})^{٢}}{\text{ن}} \right]$$

$$\therefore \text{ب} = \frac{\text{مجـ س ص} - \frac{\text{مجـ س . مجـ ص}}{\text{ن}}}{\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{\text{ن}}}$$

وتسمى ب فى هذه المعادلة بمعامل خط انحدار ص/س .

ويلاحظ على معادلتى أ ، ب وجود الكميات الاحصائية مجـ ص ، مجـ س ، مجـ س ص ، ... وأنه للحصول على هذه الكميات يستلزم تكوين الجـدول

(*) تسمى المعادله (١) ، المعادلة (٢) بالمعادلتان الطبيعيتان ويحلهما جبريا فـى أن واحد نحصل على قيمة أ ب .



الاحصائى اللازم .

تكوين الجدول الاحصائى اللازم (للمثال) :

س	ص	س ص	مجـ $س^2$	مجـ $ص^2$
٢	٢	٤	٤	٤
٣	٤	١٢	٩	١٦
٥	٦	٣٠	٢٥	٣٦
٧	٦	٤٢	٤٩	٣٦
٨	٧	٥٦	٦٤	٤٩
٢٥	٢٥	١٤٤	١٥١	١٤١

وعليه يكون مجـ س = ٢٥ ، مجـ ص = ٢٥ ،مجـ س ص = ١٤٤ ،

مجـ $س^2$ = ١٥١

$$\therefore ب = \frac{\text{مجـ س ص} - \frac{\text{مجـ س . مجـ ص}}{ن}}{\text{مجـ } س^2 - \frac{(\text{مجـ س})^2}{ن}}$$

$$\therefore ب = \frac{١٤٤ - \frac{٢٥ \times ٢٥}{٥}}{١٥١ - \frac{٢٥ \times ٢٥}{٥}}$$

$$= \frac{١٤٤ - ١٢٥}{١٥١ - ١٢٥} = \frac{١٩}{٢٦} = ٠,٧٣$$

$$، \therefore أ = \frac{\text{مجـ ص} - \text{ب مجـ س}}{ن}$$

$\therefore$ أ = $\frac{٢٥ - ٠,٧٣ \times ٢٥}{٥} = \frac{٢٥ - ١٨,٢٧}{٥} = ١,٣٥$

$\therefore$ ص̂ = ١,٣٥ + ٠,٧٣ س

وهذه هى معادلة الخط المستقيم لانحدار ص/س ، وهى ذات ميل موجب حيث معامل س ذو قيمة موجبة ، ويعنى الميل ٠,٧٣ أنه اذا تغيرت س بالزيادة بمقدار الوحدة فإن ص̂ تتغير بالزيادة بمقدار أقل من الوحدة أى بمقدار ٠,٧٣ .

٤- جدول الفروق (الإنحرافات) بين قيم ص ، ص̂ عند كل قيم س :

س	ص	ص̂= ١,٣٥+٠,٧٣س	(ص-ص̂)	(ص-ص̂)^{2}
٢	٢	ص̂=١,٣٥+٠,٧٣×٢= ٢,٨٤	-٠,٨١	٠,٦٥٦١
٣	٤	ص̂=١,٣٥+٠,٧٣×٣= ٣,٥٤	٠,٤٦	٠,٢١١٦
٥	٦	ص̂=١,٣٥+٠,٧٣×٥= ٥,٠٠	١,٠٠	١,٠٠٠٠
٧	٦	ص̂=١,٣٥+٠,٧٣×٧=٦,٤٦	-٠,٤٦	٠,٢١١٦
٨	٧	ص̂=١,٣٥+٠,٧٣×٨=٧,١٩	-٠,١٩	٠,٠٣٦١
			مجـ=صفر	مجـ=٢,١٢

ويلاحظ فى الجدول السابق ما يلى :

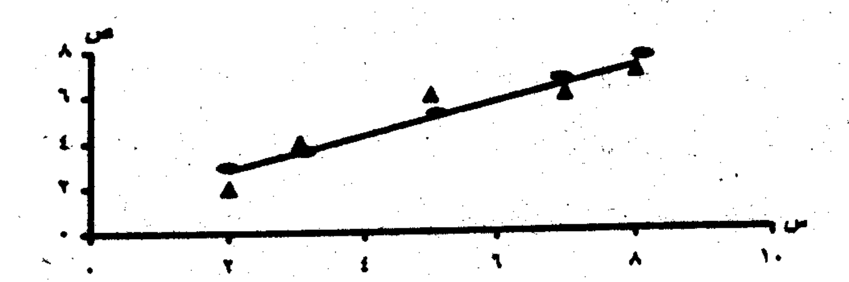
١- أن مجـ (ص-ص̂) يساوى الصفر ، وهذا يتفق مع خاصية المتوسط الحسابى بأن مجموع انحرافات القيم عن متوسطها الحسابى يساوى الصفر .

٢- أن مجـ (ص-ص̂)^{2} يساوى ٢,١٢ وهو أقل ما يمكن عما لو تم حساب هذا المجموع على اساس ص̂ لخط مستقيم آخر لانحدار



ص/س . وهذا يتفق مع خاصية المتوسط الحسابى بأن مجموع مربعات انحرافات القيم عن متوسطها الحسابى اقل ما يمكن .

ويمكن إيضاح هذه الفروق بيانيا كما فى الشكل التالى :



ويتضح من هذا الشكل نقاط شكل الانتشار (▲) ، ونقاط الخط المستقيم (▬) لانحدار ص/س ، والفروق بين ص ، $\hat{ص}$.

٥- اختبار مدى دقة التوفيق في تقدير الخط المستقيم لانحدار ص/س :

لاجراء هذا الاختبار يستلزم ايجاد :

- الخطأ المعيارى Standered Error لتقدير الخط المستقيم لانحدار ص/س أى ($خ_{ص/س}$) .
- الخطأ المعيارى لتقدير معامل الخط المستقيم لانحدار ص/س أى ($خ_ب$) .

والخطأ المعيارى ($خ_{ص/س}$) هو الجزر التربيعى لمتوسط مجموع مربعات انحرافات (فروق) قيم ص عن قيم $\hat{ص}$ المناظرة لها .

$$\therefore خ_{ص/س} = \sqrt{\frac{مج\,(ص - \hat{ص})^٢}{ن}}$$

$$\therefore خ_{ص/س} \text{ فى حالة المثال } = \sqrt{\frac{٢,١٢}{٥}} = ٠,٦٥$$

وجدير بالذكر بأن دقة التقدير تتناسب عكسيا مع مقدار الخطأ المعيارى ، فكلما صغرت قيمة الخطأ المعيارى كلما اقتربت قيم ص الفعلية من قيم صً أى كلما اقتربت قيم ص الفعلية من خط الانحدار وبالتالى ازدادت الدقة الناتجة عن استخدام معادلة الانحدار فى التفسير أو التنبؤ .

كما تجدر الاشارة الى أنه يمكن التوصل الى صورة أسهل لمعادلة $\text{خ}_{\text{ص/س}}$ كما يلى :

$$\because \text{خ}_{\text{ص/س}} = \sqrt{\frac{\text{مج}\,(\text{ص}-\hat{\text{ص}})^{2}}{\text{ن}}}$$

$$= \sqrt{\frac{\text{مج}\,(\text{ص} - \text{أ} - \text{ب س})^{2}}{\text{ن}}}$$

$$= \sqrt{\frac{\text{مج ص}^{2} - \text{أ مج ص} - \text{ب مج س ص}}{\text{ن}}}$$

وهذه هى الصورة الأكثر استخداما ، ولاثبات أن $\text{مج}\,(\text{ص}-\hat{\text{ص}})^{2}$ تساوى (مج ص^{2} – مج أ ص – ب مج س ص) فإن :

$$\therefore \text{مج}\,(\text{ص}-\hat{\text{ص}})^{2} = \text{مج}\,(\text{ص} - \text{أ} - \text{ب س})^{2} \quad \text{حيث } \hat{\text{ص}} = \text{أ} + \text{ب س}$$

$$= \text{مج}(\text{ص}^{2} - ٢\,\text{أ ص} - ٢\,\text{ب س ص} + \text{أ}^{2} + ٢\,\text{أ ب س} + \text{ب}^{2}\,\text{س}^{2})$$

$$= \text{مج ص}^{2} - ٢\text{أ مج ص} - ٢\text{ب مج س ص} + \text{ن أ}^{2} + ٢\text{أ ب مج س} + \text{ب}^{2}\,\text{مج س}^{2}$$



= مجـ ص^٢ – أ مجـ ص – ب مجـ س ص – أ مجـ ص
–ب مجـ س ص +ن أ^٢ + ٢ أ ب مجـ س + ب^٢ مجـ س^٢

= مجـ ص^٢– أ مجـ ص – ب مجـ س ص– أ (ن أ+ب مجـ س)
– ب (أمجـ س + ب مجـ س^٢) + ن أ^٢ + ٢ أ ب مجـ س
+ ب^٢ مجـ س^٢ .

=مجـ ص^٢– أ مجـ ص– ب مجـ س ص– ن أ^٢ – ب أ مجـ س
– أ ب مجـ س – ب^٢ مجـ س^٢ + ن أ^٢ + ٢ أ ب مجـ س
+ ب^٢ مجـ س^٢

= مجـ ص^٢ – أ مجـ ص – ب مجـ س ص

وهو المطلوب اثباته

وبتطبيق الصورة الأخيرة (الأكثر استخداما) على المثال فإن :

∴ خص/س = $\sqrt{\frac{\text{مجـ ص}^2 - \text{أ مجـ ص} - \text{ب مجـ س ص}}{\text{ن}}}$

= $\sqrt{\frac{١٤١ - ١,٣٥ \times ٢٥ - ٠,٧٣ \times ١٤٤}{٥}}$

= $\sqrt{\frac{٢,١٣}{٥}}$ = ٠,٦٥

وهى نفس النتيجة السابقة إلا أن الصورة الأخيرة لمعادلــــة خص/س هـــى الأكثر استخداما .



- الخطأ المعيارى لمعامل الخط المستقيم لانحدار ص/س ($خ_ب$) :

هو الخطأ المعيارى ص/س بالنسبه لمجموع مربعات انحرفات س عن $\bar{س}$

$$\therefore خ_ب = \sqrt{\frac{خ_{ص/س}}{مجـ (س-\bar{س})^2}}$$

$$= \frac{خ_{ص/س}}{\sqrt{مجـ س^2 - \frac{(مجـ س)^2}{ن}}}$$

والمعادلة الأخيرة هى الاسهل فى العمليات الحسابية ومن ثم هى الاكثر استخداما .

$$\therefore خ_ب \text{ للمثال} = \frac{0{,}65}{\sqrt{151 - \frac{25 \times 25}{5}}} = \frac{0{,}65}{\sqrt{26}} = 0{,}127$$

وإذا كانت $خ_ب$ أقل من نصف معامل الانحدار (ب) ، فإنه يوجد احتمال كبير أن يكون معامل الانحدار (ب) معنوى أى حقيقى ومن ثم فالمعادلة الناتجة يمكن الاطمئنان إليها فى التفسير أو التنبؤ .

ولكى نصل إلى مرحلة التأكد من معنوية معامل الانحدار (ب) ، فإن ذلك يتطلب إجراء اختبار (ت) وهو أن $ت = \frac{ب}{خ_ب}$ وهذا يتوقف على توزيع (ت) وهو من التوزيعات الاحتمالية والتى سيتم تناولها فى الجزء الثانى من هذا الكتاب إنشاء الله .

وتفسير $خ_ب$ الناتجة من حل المثال التى تساوى ٠,١٢٧ ومقارنتها بمعامل الانحدار (ب) والذى يساوى ٠,٧٣ فى معادلة الخط المستقيم لانحدار ص/س والتى تساوى $\hat{ص}$ = ١,٣٥ + ٠,٧٣ س نجد أن $خ_ب$ أقل من نصف معامل س وعليه فإن ب معنوية .



نتائج هامه على الانحدار البسيط ذو الخط المستقيم :

[١] $$ب = ر \times \frac{ع_{ص}}{ع_{س}}$$

حيث :

ب : هى معامل انحدار ص/س

ر : معامل الارتباط البسيط ذو الخط المستقيم للمتغيرين س،ص

$ع_{ص}$: الانحراف المعيارى للمتغير ص

$ع_{س}$: الانحراف المعيارى للمتغير س

الاثبـــات

$$\therefore ب = \frac{\text{مجـ س ص} - \frac{\text{مجـ س . مجـ ص}}{ن}}{\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{ن}}$$

$$= \frac{\text{مجـ س ص} - \frac{\text{مجـ س.مجـ ص}}{ن}}{\sqrt{\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{ن}} \times \sqrt{\text{مجـ س}^{٢} - \frac{\text{مجـ س}^{٢}}{ن}}} \times \frac{\sqrt{\text{مجـ ص}^{٢} - \frac{(\text{مجـ ص})^{٢}}{ن}}}{\sqrt{\text{مجـ ص}^{٢} - \frac{(\text{مجـ ص})^{٢}}{ن}}}$$

$$= ر \times \frac{ع_{ص}}{ع_{س}}$$

ويفيد هذا القانون فى ايجاد معامل انحدار ص/س بمعلومية معامل الارتباط (ر) والانحراف المعيارى للمتغيرين ص،س .

٢ $$\text{ر}^{2} = 1 - \frac{\text{خ}^{2}_{\text{ص/س}}}{\text{ع}^{2}_{\text{ص}}}$$

حيث :

ر^{2} : مربع معامل الارتباط البسيط ذو الخط المستقيم

$\text{خ}^{2}_{\text{ص/س}}$: مربع الخطأ المعيارى لتقدير خط المستقيم لانحدار ص/س

$\text{ع}^{2}_{\text{ص}}$: تباين المتغير ص .

الاثبـــات

$$\therefore \text{ب} = \text{ر} \cdot \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}}$$

$$، \therefore \text{أ} = \bar{\text{ص}} - \text{ب}\,\bar{\text{س}} = \bar{\text{ص}} - \text{ر} \cdot \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}} \cdot \bar{\text{س}}$$

$$، \therefore \text{خ}^{2}_{\text{ص/س}} = \frac{\text{مج ص}^{2} - \text{أ مج ص} - \text{ب مج س ص}}{\text{ن}^{2}}$$

$$\therefore \text{خ}^{2}_{\text{ص/س}} = \frac{1}{\text{ن}^{2}} \Big[\text{مج ص}^{2} - \text{مج ص} \left(\bar{\text{ص}} - \text{ر} \cdot \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}} \cdot \bar{\text{س}} \right) - \text{مج س ص} \left(\text{ر} \cdot \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}} \right) \Big]$$

$$= \frac{1}{\text{ن}^{2}} \Big[\text{مج ص}^{2} - \frac{(\text{مج ص})^{2}}{\text{ن}} + \text{ر} \cdot \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}} \cdot \frac{\text{مج ص} \cdot \text{مج س}}{\text{ن}} - \text{ر} \cdot \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}} \cdot \text{مج س ص} \Big]$$

$$= \frac{1}{\text{ن}^{2}}\left[\text{ن}^{2}\,\text{ع}^{2}_{\text{ص}} + \text{ر}\,.\,\frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}}\left(\frac{\text{مجـ ص}\,.\,\text{مجـ س}}{\text{ن}} - \text{مجـ س ص}\right)\right]$$

$$= \frac{1}{\text{ن}^{2}}\left[\text{ن}^{2}\,\text{ع}^{2}_{\text{ص}} - \text{ر}\,.\,\frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}}\left(\text{مجـ س ص} - \frac{\text{مجـ ص}\,.\,\text{مجـ س}}{\text{ن}}\right)\right]$$

$$= \frac{1}{\text{ن}^{2}}\left[\text{ن}^{2}\,\text{ع}^{2}_{\text{ص}} - \text{ر}\,.\,\frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}}\left(\text{ر}\,.\,\text{ن}^{2}\,\text{ع}_{\text{ص}}\,.\,\text{ع}_{\text{س}}\right)\right]$$

$$= \frac{1}{\text{ن}^{2}}\left[\text{ن}^{2}\,\text{ع}^{2}_{\text{ص}} - \text{ر}^{2}\,\text{ن}^{2}\,\text{ع}^{2}_{\text{ص}}\right]$$

$$= \frac{1}{\text{ن}^{2}}\,.\,\text{ن}^{2}\,\text{ع}^{2}_{\text{ص}}\,(1 - \text{ر}^{2})$$

$$= \text{ع}^{2}_{\text{ص}}\,(1 - \text{ر}^{2})$$

$$\therefore\ 1 - \text{ر}^{2} = \frac{\text{خ}^{2}_{\text{ص/س}}}{\text{ع}^{2}_{\text{ص}}}$$

$$\therefore\ \text{ر}^{2} = 1 - \frac{\text{خ}^{2}_{\text{ص/س}}}{\text{ع}^{2}_{\text{ص}}}$$

وهو المطلوب اثباته ...

٣ $(\text{ص} - \bar{\text{ص}}) = \text{ب}\,(\text{س} - \bar{\text{س}})$

وهذه معادلة خط مستقيم لانحدار ص/س لكن يمكن ايجادها بمعلومية كلا من ص ، س وبمعلومية ب سواء مباشرة أو بمعلومية ر ، $\text{ع}_{\text{ص}}$ ، $\text{ع}_{\text{س}}$.

الاثبـــات

$\therefore$ ص = أ + ب س

، $\therefore$ أ = $\bar{ص}$ - ب $\bar{س}$

$$= \bar{ص} - ر \times \frac{ع_{ص}}{ع_{س}} \times \bar{س}$$

، $\therefore$ ب $= ر \times \frac{ع_{ص}}{ع_{س}}$

$$\therefore ص = \bar{ص} - ر \times \frac{ع_{ص}}{ع_{س}} \times \bar{س} + ر \times \frac{ع_{ص}}{ع_{س}} \times س$$

$$= \bar{ص} - ر \times \frac{ع_{ص}}{ع_{س}} \times (س - \bar{س})$$

$\therefore$ (ص - $\bar{ص}$) = ب (س - $\bar{س}$)

فإذا فرض أن المعلومات المتاحة عن تغـــيرين س،ص هـــى أن المتوسط الحسابى ($\bar{س}$) للمتغير س هو ٧ وللمتغير ص ١٢ وأن معـــامل الارتباط بينها ٠,٨ وأن الانحراف المعيارى للمتغير س هو ١,٦ وللمتغير ص ١,٢ فأوجد معادلة الانحدار المستقيم ص/س .

الاجابـــة

$$\therefore (ص - \bar{ص}) = ر \times \frac{ع_{ص}}{ع_{س}} (س - \bar{س})$$

$$ص - ١٢ = ٠,٨ \times \frac{١,٢}{١,٦} (س - ٧)$$

ص = ٠,٦ س + ٧,٨ وهى المعادلة المطلوبة ..

٤ ب × بَ = $ر^٢$

حيث :-

ب : معامل معادلة انحدار ص/س

بَ : معامل معادلة انحدار س/ص

$ر^٢$: مربع معامل الارتباط البسيط للمتغيرين س ، ص .

مثـــال

نطبق هذا القانون على بيانات المثال السابق

سبق ايجاد معامل انحدار ص/س وتبين أن ب = ٠,٧٣ ولإيجاد قيمـة بَ أى معامل انحدار س/ص فيتم الآتى :-

تكوين الجدول الاحصائى اللازم :

ص	س	ص س	$ص^٢$
٢	٢	٤	٤
٤	٣	١٢	١٦
٦	٥	٣٠	٣٦
٦	٧	٤٢	٣٦
٧	٨	٥٦	٤٩
٢٥	٢٥	١٤٤	١٤١

$$ب = \frac{مجـ س ص - \frac{مجـ س . مجـ ص}{ن}}{مجـ ص^{٢} - \frac{(مجـ ص)^{٢}}{ن}}$$

$$= \frac{١٩}{١٤١ - \frac{٢٥ \times ٢٥}{٥}} = \frac{١٩}{١٦} = ١,١٩$$

$$، \therefore ر = \frac{مجـ س ص - \frac{مجـ س . مجـ ص}{ن}}{\sqrt{\left(مجـ س^{٢} - \frac{(مجـ س)^{٢}}{ن}\right)\left(مجـ ص^{٢} - \frac{(مجـ ص)^{٢}}{ن}\right)}}$$

$$= \frac{١٩}{\sqrt{٢٦ \times ١٦}} = ٠,٩٣$$

$\therefore ر^{٢} = ٠,٨٦$

$، \therefore ب \times \overline{ب} = ٠,٧٣ \times ١,١٩$

$= ٠,٨٦$

$\therefore ب \times \overline{ب} = ر^{٢}$

ولايجاد معادلة انحدار س/ص فالمتبقى هو ايجاد قيمة أ .

$$\therefore أ = \frac{مجـ س - ب مجـ ص}{ن}$$

$$= \frac{٢٥ - ١,١٩ \times ٢٥}{٥} = - ٠,٩٥$$

$\therefore \hat{ش} = أ + ب ص$



.. ش = - ٠,٩٥ + ١,١٩ ص وهى معادلة انحدار س/ص .

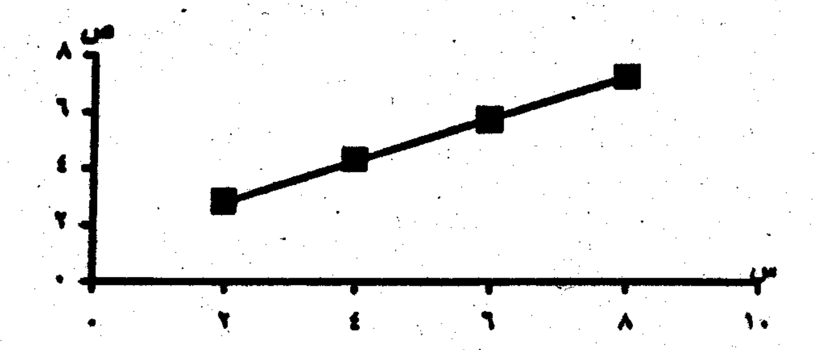
[٥] يتقاطع خطى الانحدار ص/س ، س/ص فى نقطة احداثياها السينى والصادى هما (سَ ، صَ) .

الاثبـــات

العرض الجدولي والعرض البياني لخط انحدار ص/س الناتج وهو :

ص = ١,٣٥ + ٠,٧٣ س

س	٢	٤	٦	٧
ص	٢,٨	٤,٣	٥,٧	٧,٢

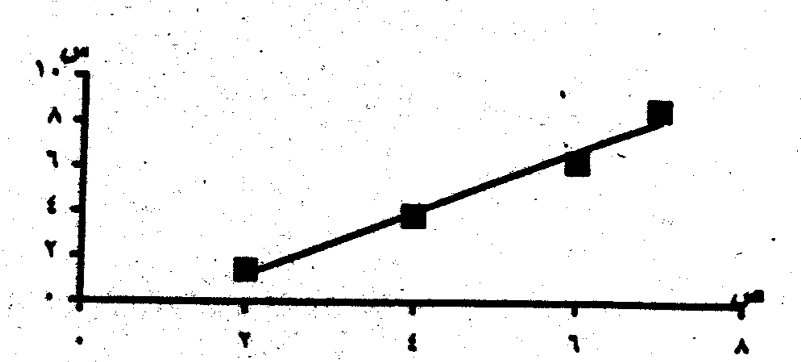


وهذا الخط يعبر عن العلاقة بين قيم المتغير س ومتوسطات القيم العاديـــة التي تناظرها .

العرض الجدولي و العرض البياني لخط انحدار س/ص الناتج وهو :

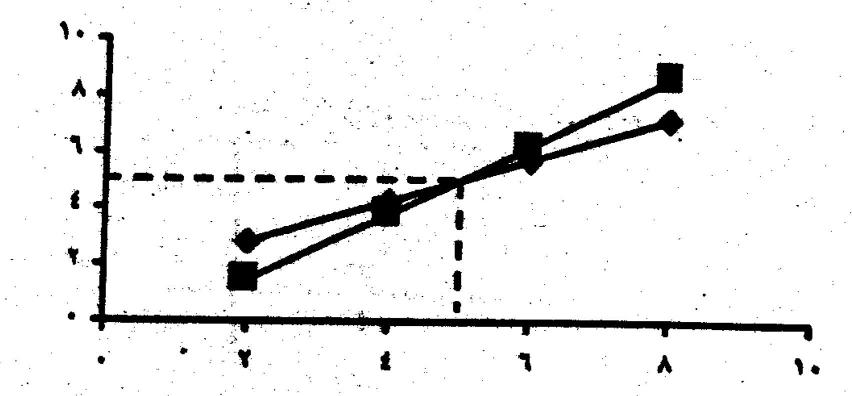
س = - ٠,٩٥ + ١,١٩ ص

س	٢	٤	٦	٧
ص	١,٤٣	٣,٨	٦,٢	٨,٦



وهذا الخط يعبر عن العلاقة بين قيم المتغير ص ومتوسطات القيم السينية التى تناظرها .

العرض البيانى لخطى انحدار ص/س ، س/ص



ويتضح من هذا العرض البيانى أن نقطة تقاطع خطـــى الانحـــدار هـــى (٥ ، ٥) ويمكن اثبات ذلك جبريا كما يلى :

.. (ص − صَ) = ب (س − سَ)

١,٣٥ + ٠,٧٣ س − صَ = ٠,٧٣ (س − سَ)

.. صَ − ٠,٧٣ سَ = ١,٣٥ .... (١)

$\therefore$ (س – $\bar{س}$) = ب (ص – $\bar{ص}$)

– ٠,٩٥ + ١,١٩ ص – س = ١,١٩ (ص – $\bar{ص}$)

$\therefore$ $\bar{س}$ – ١,١٩ $\bar{ص}$ = – ٠,٩٥ .... (٢)

وبحل المعادلتين (١)،(٢) جبريا ينتج أن :

النقطة (س ، ص) = (٥ ، ٥)

٦ خط الانحدار البسيط ذو الخط المستقيم مـــن الجـدول التكــرارى المــزدوج:

مثال

الجدول التالى هو جدول توزيع تكرارى مزدوج لمتغيرين س ، ص

فئات س / فئات ص	(٠–٢)	(٢–٤)	(٤–٦)	(٦–٨)	مجموع التكرارات
(١ – ٣)	١	١			٢
(٣ – ٥)		٢	١		٣
(٥ – ٧)			٢	١	٤
(٧ – ٩)			١	٢	٣
مجموع التكرارات	١	٣	٤	٣	١٢

والمطلوب :

احسب خط انحدار ص/س

الاجابة

بداية نوضح أنه في حالة البيانات التي تم تجميعها عن المتغيرات س ، ص من مجتمع احصائي صغير (بيانات غير مبوبة) كنا نلاحظ أن كل قيمة من قيم المتغير س تناظر قيمة من قيم المتغير ص ، لكن في حالة المجتمع الاحصائي الكبير (بيانات مبوبة) واضطرارنا إلى عمل جدول توزيع تكراري مزدوج فنلاحظ ظهور المتغير س على شكل فئات ذات تكرارات مرتبطة بفئات ص المناظرة لها ، وعليه فالحساب خط انحدار ص/س في هذه الحالة يتم الآتي :

* تكوين الجدول الاحصائي اللازم :

فئات س	١	٣	٥	٧
ص المناظرة لفئات س المختلفة	٢	٤	٦	٧,٣

بديهي معرفة أن مركز الفئة س يتأتى من $\frac{\text{الحد الأعلى + الحد الادنى}}{٢}$

- أما ص المناظرة لفئات س المختلفة تتأتى من

ص المناظرة لأى فئة س = $\frac{\text{مج (مركز فئات ص × التكرار المناظر لها تحت الفئة س)}}{\text{مج تكرار الفئة س}}$

فمثلا ص المناظرة للفئة (٢-٤) = $\frac{\text{مج} (٢×١ + ٤×٢ + ٦×١)}{٤}$ = ٤

، ص المناظرة للفئة (٦-٨) = $\frac{\text{مج} (٦×١ + ٨×٢)}{٣}$ = ٧,٣٠



• تكوين الجدول الاحصائى اللازم ② :

س	ص	س ص	$\text{س}^{٢}$
١	٢	٢	١
٣	٤	١٢	٩
٥	٦	٣٠	٣٥
٧	٧,٣	٥١,١	٤٩
١٦	١٩,٣	٩٥,١	٨٤

$$\therefore \text{ب} = \frac{\text{مجـ س ص} - \frac{\text{مجـ س . مجـ ص}}{\text{ن}}}{\text{مجـ س}^{٢} - \frac{(\text{مجـ س})^{٢}}{\text{ن}}}$$

$$\therefore \text{ب} = \frac{٩٩,١ - \frac{١٦ \times ١٩,٣}{٤}}{٨٤ - \frac{١٦ \times ١٦}{٤}} = \frac{٩٥,١ - ٧٧,٢}{٨٤ - ٦٤} = ٠,٩$$

$$، \therefore \text{أ} = \frac{\text{مجـ ص} - \text{ب مجـ س}}{\text{ن}} = \frac{١٩,٣ \times ٠,٩ \times ١٦}{٤} = ١,٢$$

$\therefore \hat{\text{ص}} = ١,٢ + ٠,٩ \text{ س}$ وهى المعادلة المطلوبة .

الانحدار البسيط ذو الخط المنحنى :

فى بعض الأحيان نجد أن شكل الانتشار يوضح أن العلاقة بين المتغيرين س ، ص لا يمكن تقريب نقاطها بخط مستقيم مما يعنى أن التقريب الممكن هو باستخدام خط منحنى .



أولا : إذا كان الخط المنحنى يتفق مع معادلة من الدرجة الثانية :

ص = أ + ب س + جـ $س^2$

مثـــال

الجدول التالى يوضح بيانات عن قيم المتغير س وقيم المتغير ص المناظرة لها :

س	٢	٥	٨	١١	١٤
ص	٢	١٠	١٩	١٦	١٨

والمطلوب :

١-ارسم شكل الانتشار

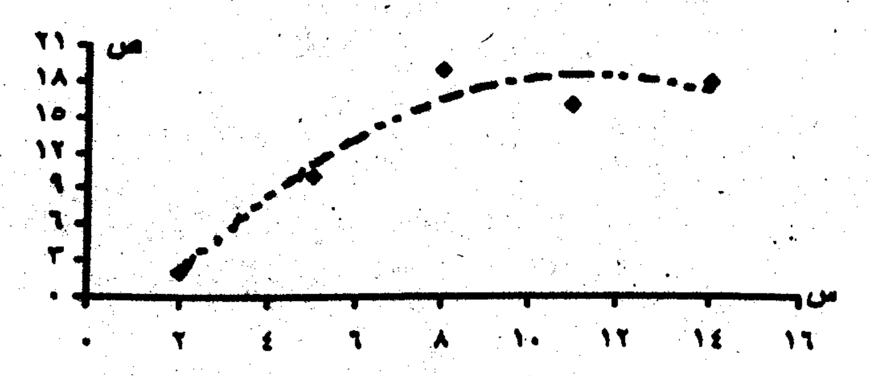
٢-أوجد خط انحدار ص/س على فرض أنه معادلة من الدرجة الثانية مستخدما طريقة المربعات الصغرى .

٣- اختبر اختيارك للمعادلة السابقة كمعادلة تحكم سلوك المتغيرين س ، ص .

٤-كيف يمكن ايجاد معامل الارتباط فى هذه الحالة .

الاجابـــة

١-شكل الانتشار





هذا هو شكل الانتشار موضحا به خط انحـــدار ص/س بمجـــرد النظر وواضح أنه يأخذ شكل منحنى من الدرجة الثانية .

٢-لايجاد خط انحدار ص/س جبريا على فرض أنه معادلة من الدرجـــة الثانية ص= أ + ب س + جـــ س^{٢} بطريقة المربعات الصغرى نتبـــع الآتي :-

تكوين الجدول الاحصائي اللازم :

س	ص	س ص	س^{٢}	س^{٢}ص	س^{٣}	س^{٤}	ص^{٢}
٢	٢	٤	٤	٨	٨	١٦	٤
٥	١٠	٥٠	٢٥	٢٥٠	١٢٥	٦٢٥	١٠٠
٨	١٩	١٥٢	٦٤	١٢١٦	٥١٢	٤٠٩٦	٣٦١
١١	١٦	١٧٦	١٢١	١٩٣٦	١٣٣١	١٤٦٤١	٢٥٦
١٤	٨	١١٢	١٩٦	١٥٦٨	٢٧٤٤	٣٨٤١٦	٦٤
٤٠	٥٥	٤٩٤	٤١٠	٤٩٧٨	٤٧٢٠	٥٧٧٩٤	٧٨٥

وباستخدام طريقة المربعات الصغرى في التوصل الى أفضل خط منحنـــى لانحدار ص/س أي صَ = أ + ب س + جـــ س^{٢} تمثل بيانـــــات شـــكل الانتشار ، فإن الأمر يتطلـــب إيجـــاد مجـــاهيل تلـــك المعادلـــة وهـــى أ ، ب ، جـــ كما يلي• :

∴ مجـ ف = مجـ (صَ - ص)^{٢}

= مجـ (أ + ب س + جـــ س^{٢} - ص)^{٢}

∴ $\frac{د ف}{د أ}$ = ٢ مجـ (أ + ب س + جـــ س^{٢} - ص) ومساوتها بالصفر

• يستخدم التفاضل الأول للدالة بالنسبة لـــ أ ، ب ، جـــ



∴ مجـ ص = ن أ + ب مجـ س + جـ مجـ س^{٢} (١)

∴ $\frac{د ف}{د ب}$ = ٢ مجـ (أ + ب س + جـ س - ص) × س ومساوتها بالصفر

∴ مجـ س ص = أ مجـ س + ب مجـ س^{٢} + جـ مجـ س^{٣} (٢)

∴ $\frac{د ف}{د جـ}$ = ٢مجـ (أ + ب س + جـ س^{٢} - ص) × س^{٢} مساواتها بالصفر

∴ مجـ س^{٢} ص = أ مجـ س^{٢} + ب مجـ س^{٣} + جـ مجـ س^{٤} (٣)

وتسمى المعادلات الثلاث الناتجة بالمعادلات الطبيعية ، وبحلها جبريا في آن واحد نحصل على قيم أ ، ب ، جـ ومن ثم الحصول على المعادلة المطلوبة :

٥٥ = ٥أ + ٤٠ب + ٤١٠ جـ •
٤٩٤ = ٤٠أ + ٤١٠ب + ٤٧٢٠ جـ
٤٩٧٨ = ٤١٠أ + ٤٧٢٠ب + ٥٧٧٩٤ جـ

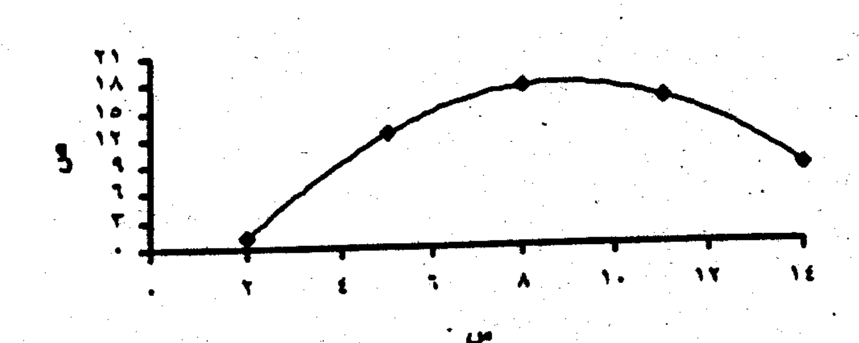
∴ أ = - ٩,٩ ، ب = ٦.٢ ، جـ = - ٠,٣٥

∴ ص^ = - ٩.٩ + ٦,٢ س - ٠,٣٥ س^{٢} وهي المعادلة المطلوبة

و يتضح هذا الخط المنحني بيانيا كما يلي :

س	٢	٥	٨	١١	١٤
ص^	١.١	١٢,٣٥	١٧.٣	١٥,٩٥	٨,٣

• بالتعويض بقيمة أ = ١١ - ٨ ب - ٨٢ جـ في المعادلة الثانية والثالثة .



وهذا الخط المنحنى يعبر عن العلاقة بين قيم المتغير س ومتوسطات القيم الصادية التى تناظرها .

٤- اختبار مدى دقة التوفيق فى تقدير الخط المنحنى من الدرجة الثانية لانحدار ص/س .

$$\therefore \text{خ}_{\text{ص/س}} = \sqrt{\frac{\text{مجـ ص}^2 - \text{أ مجـ ص} - \text{ب مجـ س ص} - \text{جـ مجـ س}^2\text{ ص}}{\text{ن}}}$$

$$\therefore \text{خ}_{\text{ص/س}} = \sqrt{\frac{٧٨٥ - (-٩,٩ \times ٥٥) - ٤٩٤ \times ٦,٢ - (-٠,٣٥ \times ٤٩٧٨)}{٥}}$$

$$= \sqrt{\frac{٩}{٥}} = ١,٣٤$$

٤- قياس الارتباط فى المثال الحالى محل الدراسة :

$\therefore \text{ر}^2 = ١ - \frac{\text{خ}^2_{\text{ص/س}}}{\text{ع}^2_{\text{ص}}}$ فى حالة انحدار ص/س ذو الخط المستقيم

$\therefore$ خ$^2_{\text{ص/س}}$ فى هذا المثال هو لمنحنى وليس لخط مستقيم

---

* يمكن اثبات هذا القانون كما فى حالة خ$_{\text{ص/س}}$ فى الانحدار البسيط ذو الخط المستقيم



∴ يتم استخدام $ط^2$ بدلا من $ر^2$ لتعبر عن الارتباط فى هذه الحالة وتسمى $ط^2$ بدليل الارتباط Correlation Index

$$∴ ط^2 = ١ - \frac{خ^2_{ص/س}}{ع^2_{ص}}$$ لانحدار ص/س ذو الخط المنحنى

وعليـــه

$$∴ خ^2_{ص/س} \text{ للمثال} = \frac{٩}{٥} = ١,٨$$

$$، ∴ ع^2_{ص} = \frac{١}{ن} \left(مج ص^2 - \frac{(مج ص)^2}{ن} \right)$$

$$= \frac{١}{٥} \left(٧٨٥ - \frac{٥٥ \times ٥٥}{٥} \right)$$

$$= ٣٦$$

$$∴ ط^2 = ١ - \frac{١,٨}{٣٦}$$

$$= ٠,٩٥$$

∴ ط = ٠,٩٧ وهو ارتباط قوى جدا بين المتغـــيرين س ، ص محل الدراسة .

وجدير بالذكر أن اشارة ط لا معنى لها لأن الارتباط فـــى هـــذه الحالـــة لمنحنى حيث تعدد الميول للخط المنحنى وليس ميل واحد كما فى حالـــة الخط المستقيم ، ولكن بصفة عامة فى مثالنا هذا يمكن القـــول أنـــه فـــى الجزء الأول من المنحنى كان الارتباط قوى موجب بينما فى الجزء الثانى من هذا المنحنى فالارتباط قوى سالب .



ثانيا : اذا كان الخط المنحنى يتفق مع معادلة هندسية

ص = أ س ب

مثـــال

الجدول التالى يبين الدخل والانفاق لعدد ٣٠ فرد من العاملين فـى النشاط السياحى :

الدخل	٣٥,١	٣٥,٣	٣٥,٥	٣٧,٨	٣٧,٨	٤١,٤	٤١,٧	٤٤,٨	٤٤,٩
الانفاق	٢٤,٣	٢٤,١	٢٤,٤	٢٧,٨	٢٨,٧	٤٢,٠	٣٤,٥	٥٦,١	٤٩,٠

٤٧,٩	٥١,٥	٥٣,٩	٥٦,٩	٥٨,١	٦٠,٠	٦٣,٠	٦٦,٢	٦٧,١	٦٩.٢
٥٨,٤	٦٣,٤	٧٨,٧	٨٠,٧	٨٠,٠	٩٦,٢	١١٢,٨	١١٥,٣	١٢٥,٦	١٤٦,٦

٦٩,٥	٧٠,٧	٧٣,١	٧٣,١	٧٧,٤	٨١,٧	٨١,٧	٨٢,٣	٨٣,١
١٣٢,٦	١٤٢,٨	١٣٧,١	١٦٣,٢	١٨٠,٠	١٩٨,٠	٢٠٧,٨	١٩٠,٨	٢٢٥,٥

٨٤,٦	٩٢,٤
٢٣٧,٠	٣٠٠,٢

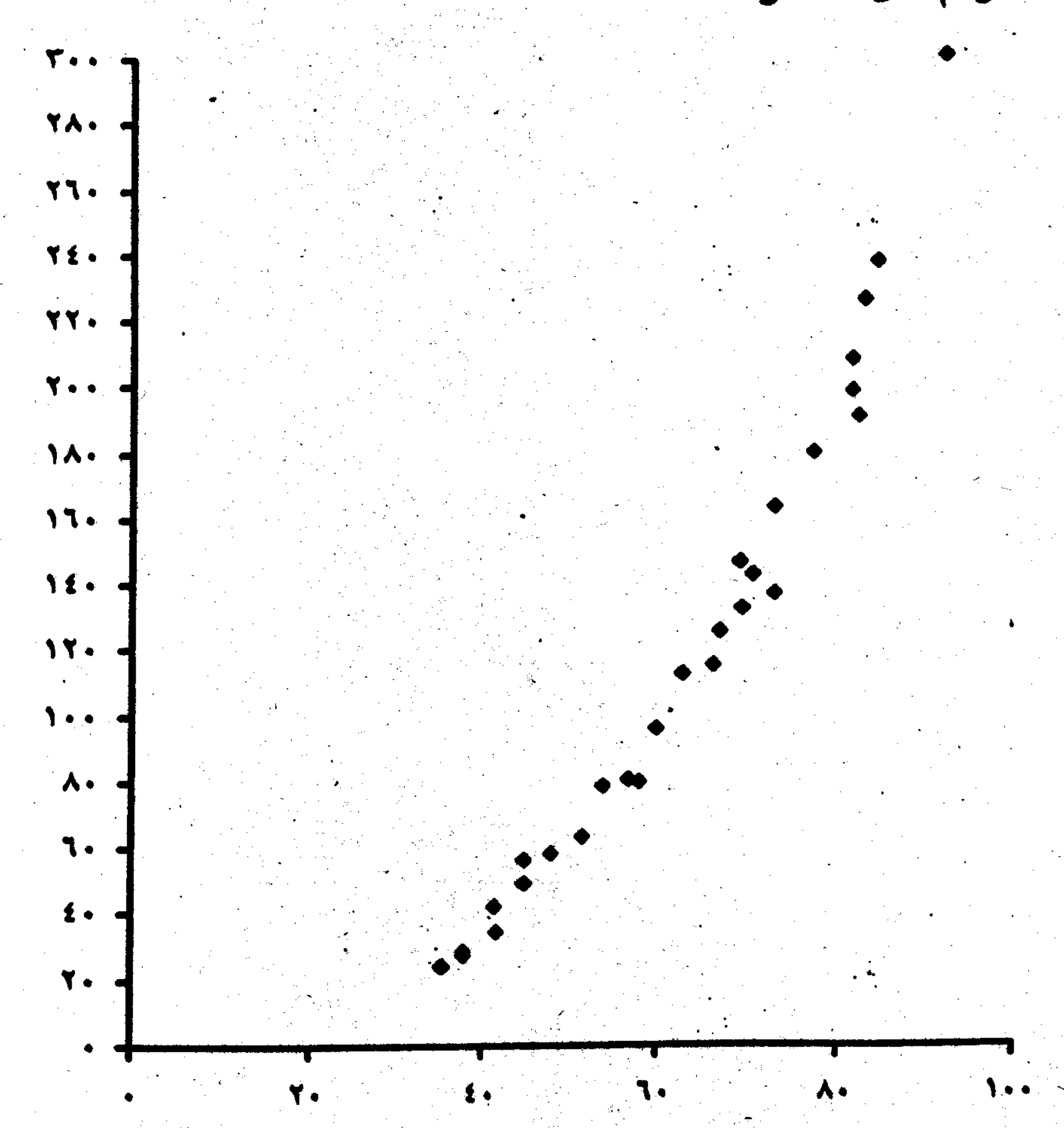
والمطلوب :

١-ارسم شكل الانتشار

٢-وفق أفضل خط لإنحدار ص/س

الاجابـــة

١-رسم شكل الانتشار :



٢- توفيق أفضل خط انحدار ص/س :

يوضح شكل الانتشار وجود نموذج واضح لاتجاه انتشار النقط وهو خـط منحنى يخضع للمعادلة ص = أ س^ب أى علاقة قوى وأنه لتوفيق افضـل خط انحدار ص/س فى هذه الحالة ، فإنه يتم تحويل هذه المعادلـة إلـى

معادلة لوغاريتمية وهى معادلة ذات خط مستقيم ، وعليه فيمكننا استخدام طريقة المربعات الصغرى على المتغيرين فى صورتيها اللوغاريتمية فنحصل على أفضل خط مستقيم يربط بينها ، كما يمكن بعد ذلك الى الحصول على العلاقة الرياضية الاصلية بين المتغيرين الأصليين وذلك بارجاع المعادلة المقدرة من صورتها اللوغاريتمية إلى صورتها الأصلية.

وعليه

.. ص = أ س^ب

.. لو ص = لو أ + ب لو س

$$.. ب = \frac{مجـ لو س لو ص - \frac{مجـ لوس . مجـ لوص}{ن}}{مجـ لو س^{2} - \frac{(مجـ لوس)^{2}}{ن}}$$

$$لو أ = \frac{مجـ لوص - ب مجـ لوس}{ن}$$

تكوين الجدول الاحصائى اللازم :

س	ص	لوس	لوص	لوس لوص	(لوس)^٢	(لوص)^٢
٣٥,١	٢٤,٣	١,٥٥	١,٣٩	٢,١٥	٢,٢٥	١,٩٣
٣٥,٣	٢٤,١	١,٥٥	١,٥٨	٢,١٤	٢,٢٥	١,٩٠
٣٥,٥	٢٤,٤	١,٥٥	١,٣٩	٢,١٥	٢,٢٥	١,٩٢
٣٧,٨	٢٧,٨	١,٥٨	١,٤٤	٢,٢٨	٢,٥٠	٢,٠٧
٣٧,٨	٢٨,٧	١,٥٨	١,٤٦	٢,٣١	٢,٥٠	٢,١٣



تابع الجدول

س	ص	لوس	لوص	لوس لوص	(لوس)^2	(لوص)^2
٤١,٤	٤٢,٥	١,٦٢	١,٦٢	٢,٦٢	٢,٦٢	٢,٦٢
٤١,٧	٣٤,٥	١,٦٢	١,٥٤	٢,٤٩	٢,٦٢	٢,٣٧
٤٤,٨	٥٦,١	١,٦٥	١,٧٥	٢,٨٩	٢,٧٢	٣,٠٦
٤٤,٨	٤٩,٠	١,٦٥	١,٦٩	٢,٧٩	٢,٧٢	٢,٨٦
٤٧,٩	٥٨,٤	١,٦٥	١,٧٧	٢,٩٢	٢,٧٢	٣,١٣
٥١,٥	٦٣,٤	١,٧١	١,٨٠	٣,٠٨	٢,٩٢	٣,٢٤
٥٣,٩	٧٨,٧	١,٧٣	١,٩٠	٣,٢٩	٣,٠٠	٣,٦١
٥٦,٩	٨٠,٧	١,٧٦	١,٩١	٣ ٣٦	٣,١	٣,٦٥
٥٨,١	٨٠,٠	١,٧٦	١,٩٠	٣.٣٤	٣,١	٣,٦١
٦٠,٠	٩٦,٢	١,٧٨	١,٩٨	٣,٥٢	٣,١٧	٣,٩٢
٦٣,٠	١١٢,٨	١,٨٠	٢,٠٥	٣,٦٩	٣,٢٤	٤,٢٠
٦٦,٢	١١٥,٣	١,٨٢	٢,٠٦	٣,٧٥	٣,٣١	٤,٢٤
٦٧,١	١٢٥,٦	١,٨٣	٢,١٠	٣,٨٤	٣,٣٥	٤,٤١
٦٩,٢	١٤٦,٦	١,٨٤	٢,١٧	٣,٩٩	٣,٣٩	٤,٧١
٦٩,٥	١٣٢,٦	١,٨٤	٢,١٢	٣,٩٠	٣,٣٩	٤,٤٩
٧٠,٧	١٤٢,٨	١,٨٢	٢,١٥	٣,٩٨	٣,٤٢	٤,٦٢
٧٣,١	١٣٧,١	١,٨٦	٢,١٤	٣,٩٨	٣,٤٦	٤,٥٨
٧٣,١	١٦٣,٢	١,٨٦	٢,٢١	٤,١١	٣,٤٦	٤,٨٨
٧٧,٤	١٨٠,٠	١,٨٩	٢,٢٦	٤,٢٧	٣,٥٧	٥,١١
٨١,٧	١٩٨,٠	١,٩١	٢,٣٠	٤,٣٩	٣,٦٥	٥,٢٩
٨١,٧	٢٠٧,٨	١,٩١	٢,٣٢	٤,٤٣	٣,٦٥	٥,٣٨
٨٢,٣	١٩٠٠,٨	١,٩٢	٢,٢٨	٤,٣٨	٣,٦٩	٥,٢٠



تابع الجدول

س	ص	لوس	لوص	لوس لوص	(لوس)^٢	(لوص)^٢
٨٣,١	٢٢٥,٥	١,٩٢	٢,٣٥	٤,٥١	٣,٦٩	٥,٥٢
٨٤,٦	٢٣٧,٠	١,٩٣	٢,٣٥	٤,٥٧	٣,٧٢	٥,٦٢
٩٢,٤	٣٠٠,٢	١,٩٧	٢,٤٨	٤,٨٩	٣,٨٨	٦,١٥
١٨١٧,٢	٣٣٨٣,٦	٥٢,٩	٥٨,٢٨	١٠٤,١	٩٣,٨	١١٦,٤٣

$$\therefore \text{ب} = \frac{\text{مج لوس لوص} - \dfrac{\text{مج لوس} - \text{ب مج لوص}}{\text{ن}}}{\text{مج لوس}^{٢} - \dfrac{(\text{مج لوس})^{٢}}{\text{ن}}}$$

$$= \frac{١٠٤,١ - \dfrac{٥٨,٢٨ \times ٥٢,٩}{٣٠}}{٩٣,٣١ - \dfrac{٥٢,٩ \times ٥٢,٩}{٣٠}} = \frac{١٠٤,١ - ١٠٢,٧٧}{٩٣,٨٠ - ٩٣,٢٨}$$

$$= \frac{١,٣}{٠,٥} = ٢,٥$$

$$\text{،} \therefore \text{لوأ} = \frac{\text{مج لوص} - \text{ب مج لوس}}{\text{ن}}$$

$$= \frac{٥٨,٢٨ - ٢,٥ \times ٥٢,٩}{٣٠} = - ٢,٤٦$$

$$\therefore \hat{\text{لوص}} = - ٢,٤٦ + ٢,٥ \text{ لوس}$$

وهذه المعادلة هي العلاقة الرياضية المعبرة عن أفضل خط مستقيم يمثل العلاقة بين المتغيرين س ، ص في صورتيهما اللوغاريتمية ، وللحصول



على العلاقة الرياضية بين المتغيرين الأصليين س ، ص فإنه يتم ارجـاع هذه المعادلة من صورتها اللوغاريتمية الى صورتها الأصلية كما يلى* :

$$\hat{ص} = ٠,٠٠٣ \; س^{٢,٥}$$

ثالثا : اذا كان الخط المنحنى يتفق مع معادلة أسية :

$$ص = أ \; ب^{س}$$

تجدر الاشارة إلى أن هذه المعادلة تتناسب مع ظاهرة تطور معدلات النمو كمعدلات نمو السكان فى دولة ما .

مثال

الجدول التالى يبين تطور عدد السكان لدولة ما خــلال الفــترة الزمنيــة ١٩٠٠ – ٢٠٠٠ .

الزمن (س)	١٩٠٠	١٩٠١	٢	٣	٤	٥	٦	٧	٨	٩	٢٠٠٠
عدد السكان (ص)	٧٣١	٢٣٠٠	٢٦٣٧	١٦١٥٩	١٧٧٠٠	٣٩٥٧٨	٧٤٣٦١	١٤٧٩٩٥	٢٠٣٣٤١	٣٣٤٣٨٧	٥٧٣٢٢٤

والمطلوب :

- ارسم شكل الانتشار
- وفق : أفضل خط لانحدار ص/س

الحـل

رسم شكل الانتشار :

يتضح من الرسم البيانى التالى أنه على مساحة كبيرة من الورقة وذلـــك لأنه تم على ورق رسم بيانى عادى والمتغير التابع ذو مدى واســع فــى التغير ، ولهذا يفضل أن يكون العرض البيانى فى مثل هذه الحالات علــى

---

* إيجاد العدد المقابل للوغاريتم لطرفى المعادلة اللوغاريتمية .

ورق رسم بياني نصف لوغاريتمي ، لكن على أية حال فهذا هـــو شـــكل الانتشار .

ص

٥٧٥
٥٥٠
٥٢٥
٥٠٠
٤٧٥
٤٥٠
٤٢٥
٤٠٠
٣٧٥
٣٥٠
٣٢٥
٣٠٠
٢٧٥
٢٥٠
٢٢٥
٢٠٠
١٧٥
١٥٠
١٢٥
١٠٠
٧٥
٥٠
٢٥

٠ ١ ٢ ٣ ٤ س ٥ ٦ ٧ ٨ ٩

الباب الخامس

الدراسة الاحصائية للعلاقة بين أكثر من متغيرين أحصائيين



- توفيق أفضل خط انحدار ص/س .

يوضح شكل الانتشار وجود نموذج واضح لاتجاه انتشار النقط وهو خط منحنى يخضع للمعادلة ص = أ ب^س أي علاقة أسية :

و عليه

∴ ص = أ ب^س

∴ لو ص = لو أ + س لو ب = لو أ + لو ب س

$$\therefore \text{لو ب} = \frac{\text{مجـ س لوص} - \frac{\text{مجـ س . مجـ لوص}}{\text{ن}}}{\text{مجـ س}^2 - \frac{(\text{مجـ لوس})^2}{\text{ن}}}$$

$$\therefore \text{لو أ} = \frac{\text{مجـ لوص} - \text{ب مجـ س}}{\text{ن}}$$

تكوين الجدول الاحصائي اللازم :

السنوات	س	ص	لوص	س لوص	س^٢	(لوص)^٢
١٩٠٠	٠	٧٣١	٢,٨٦	٠	٠	٨,١٨
١٩٠١	١	٢٣٠٠	٣,٣٦	٣,٣٦	١	١١,٢٩
٢	٢	٢٦٣٧	٣,٤٢	٦,٨٤	٤	١١,٧٠
٣	٣	١٦١٥٩	٤,٢١	١٢,٦٣	٩	١٧,٧٢
٤	٤	١٧٧٠٠	٤,٢٥	١٧,٠٠	١٦	١٨,٠٦
٥	٥	٣٦٥٧٨	٤.٦٠	١٨,٠٠	٢٥	٢١,١٦
٦	٦	٧٤٣٦١	٤,٨٧	٢٩,٢٢	٣٦	٢٣,٧٢



تابع الجدول

السنوات	س	ص	لوص	س لوص	س²	(لوص)²
٧	٧	١٤٧٩٩٥	٥,١٧	٢٦,١٩	٤٩	٢٦,٧٣
٨	٨	٢٠٣٣٤١	٥,٣١	٤٢,٤٨	٦٤	٢٨,٢
٩	٩	٣٣٤٣٨٧	٥,٥٢	٤٩,٦٨	٨١	٣٠,٤٧
٢٠٠٠	١٠	٥٧٣٢٢٤	٥,٧٦	٥٧,٦٠	١٠٠	٣٣,٠٨
١١	٥٥		٤٩,٣٣	٢٧٨	٣٨٥	٢٣٠.٤١

$$\therefore \text{لو ب} = \frac{\text{مجـ س لوص} - \frac{\text{مجـ س . مجـ لوص}}{\text{ن}}}{\text{مجـ س}^{٢} - \frac{(\text{مجـ لوس})^{٢}}{\text{ن}}}$$

$$= \frac{٢٧٨ - \frac{٤٩,٣٣ \times ٥٥}{١١}}{٣٨٥ - \frac{٥٥ \times ٥٥}{١١}} = \frac{٢٧٨ - ٢٤٦,٦٥}{٣٨٥ - ٢٧٥}$$

$$= ٠,٢٩$$

$$\therefore \text{لو أ} = \frac{\text{مجـ لوص} - \text{ب مجـ س}}{\text{ن}}$$

$$= \frac{٤٩,٣٣ - ٠,٢٩ \times ٥٥}{١١} = \frac{٤٩,٣٣ - ١٥,٩٥}{١١} = ٣,٠٣$$

$$\therefore \text{لو ص̂} = ٣,٠٣ + ٠,٢٩ \text{ س}$$

$$\therefore \text{ص̂} = ١٠٧١,٥٢ \times (١,٩٥)^{\text{س}}$$

وهذه هى المعادلة التى تمثل افضل خط الانحدار ص/س .

الدراسة الإحصائية للعلاقة بين أكثر من متغيرين إحصائيين

يشتمل هذا الباب على فصلين : -

الفصل الأول : الارتباط المتعدد والارتباط الجزئى

- مفهوم الارتباط المتعدد والارتبط الجزئى
- معامل الارتباط المتعدد .
- معامل الارتباط الجزئى .

الفصل الثانى : الانحـــدار المتعـــدد

- مفهوم الانحدار المتعدد .
- الانحدار المتعدد ذو الخط المستقيم :-
  - $\hat{ص}$= د(س ، ع) = أ + ب س + جـ ع
  - $\hat{ص}$= د(س،ع،ل) =أ+ب س+جـ ع + د ل
- العلاقة بين الخطأ المعيارى للتقدير ومعنوية النمـــوذج ككل

مقدمة:

سبق دراسة تأثير ظاهرة ما بمتغيرين فقط والذى رمزنا لهما بالمتغير س والمتغير ص ، وكانت الدراسة الاحصائية لهذين المتغيرين تقوم على دراسة العلاقة بينهما (الارتباط البسيط) وعلى دراسة السببية بينهما (الانحدار البسيط).

لكن قد تتأثر ظاهرة ما بالعديد من المتغيرات مثل تأثير الكمية المطلوبة من سلعة ما بسعرها ودخل المستهلك ...، واعتماد درجة نجاح الطالب فى نهاية العام على درجة نجاحة فى الفصل الدراسى الاول وعلى درجة نجاحة فى الفصل الدراسى الثانى ، كذلك ضغط الدم لشخص ما يعتمد على وزنة وعلى عمره .

وفى هذا الباب تقوم الدراسة الاحصائية على دراسة العلاقة بين هذه المتغيرات مجتمعة (الارتباط المتعدد Multiple Correlation) ، أو دراسة العلاقة بين متغيرين فقط بافتراض ثبات العوامل الآخرى (الاتباط الجزئى Partial corrilation)، وايضا تقوم الدراسة الاحصائى فى هذا الباب على دراسة العلاقة السببية بين هذه المتغيرات (الانحدار المتعدد Multipe Regression).



الفصل الاول

الارتباط المتعدد والارتباط الجزئى

أولا: الارتباط المتعدد للعلاقة ذات الخط المستقيم:-

مثال

الجدول التالى يبين درجات الامتحان فى الاقتصاد والاحصاء والرياضة لمجموعة مكونة من ٥ طلاب بالسياحة والفنادق.

الاقتصاد (ص)	٢	٣	٥	٥	١٠
الاحصاء (س)	١	٢	٣	٤	٥
الرياضة (ى)	١	٢	١	٢	٣

والمطلوب :

احسب معامل الارتباط المتعدد بافتراض أن العلاقة بين المتغيرات الثلاثة علاقة ذات خط مستقيم.

الحـــل

نظرا لكون العلاقة بين ثلاث متغيرات فان العرض البيانى المعتاد سابقا (الديكارتى) لشكل الانتشار لا يصلح فى هذه الحالة ، وانما العرض البيانى لشكل الانتشار فى هذه الحالة هو نقاط فى الفراغ ذو الثلاث أبعاد

وهذا نادر التناول ، كما أنة اذا كنا بصدد دراسة العلاقة بين أكثر من ثلاث متغيرات فانة يستحيل دراسة هذه العلاقة بيانيا ويصبح الحل الجبرى فقط هو المستخدم فى الحالتين.

وتتعدد الطرق الجبرية فى حساب معامل الاتباط المتعدد الا أنه يفضل الطريقة التالية والتى تقوم على الخطوات الاتية:-

١- توفيق افضل خط مستقيم يمثل العلاقة بين المتغيرات الثلاث .
٢- حساب الخطأ المعيارى لخط الانحدار الناتج .
٣- حساب معامل الاتباط المتعدد .

ونتائج الحل فيما يلى

١-توفيق افضل خط مستقيم :

نفرض أن العلاقة ذات الخط المستقيم بين المتغيرات الثلاثة تخضع للمعادلة الجبرية ص = أ + ب س + ج ى ، ولايجاد افضل خط لهذه المعادلة يمثل بيانات المتغيرات الثلاثة ، فانة يتم استخدام طريقة المربعات الصغرى والتى تنتهى بايجاج قيم أ ، ب ، ج من خلال المعادلات الثلاث الطبيعية معا وهى :

مج ص = ن أ + ب مج س + جـ مج ى

مج ص س = أ مج س + ب مج س^{2} + جـ مج ى س

مج ص ى = أ مج ى + ب مج س ى + جـ مج ى^{2}



تكوين الجدول الاحصائي اللازم :

س	ص	ى	س ص	س ى	ص ى	$س^٢$	$ص^٢$	$ى^٢$
١	٢	١	٢	١	٢	١	٤	١
٢	٣	٢	٦	٤	٦	٤	٩	٤
٣	٥	١	١٥	٣	٥	٩	٢٥	١
٤	٥	٢	٢٠	٨	١٠	١٦	٢٥	٤
٥	١٠	٣	٥٠	١٥	٣٠	٢٥	١٠٠	٩
١٥	٢٥	٩	٩٣	٣١	٥٣	٥٥	١٦٣	١٩

٩٣ = ٥ أ + ١٥ ب + ٩ جـ

٩٣ = ١٥ أ + ٥٥ ب + ٣١ جـ

٩٣ = ٩ أ + ٣١ ب + ١٩ جـ

وبحل المعادلات الثلاث جبريا سواء بالحذف أو التعويض أو باســـتخدام المصفوفات فان أ = ٠,٨ ، ب = ١,٥٣ ، جـ = ٠,٦٧ .

.. ص^ = -٠,٨ ، ب = ١,٥٣ ، جـ = ٠,٦٧

، .. فى حالة المتغيرين يكون $ر^٢_{ص/س} = ١ - \frac{خ^٢_{ص/س}}{ع^٢_{ص}}$

فانة فى حالة أكثر من متغيرين وليكون ثلاثة كما فى المثال فان :

$$ر^٢_{ص/س ى} = ١ - \frac{خ^٢_{ص/س ى}}{ع^٢_{ص}}$$

$$= ١ - \frac{مج ص^٢ - أ مج ص - ب مج س ص - جـ مج ص ى}{مج ص^٢ - \frac{(مج ص)^٢}{ن}}$$



$$= ١ - \frac{١٦٣ - \times - ٠,٨ \times ٢٥ - ١,٥٣ \times ٩٣ - ٠,٦٧ \times ٥٣}{١٦٣ - \frac{٢٥ \times ٢٥}{ن}}$$

$= ٠,٨٦$

$\therefore ر_{ص/س ى} = ٠,٩٣$

أى أن الارتباط المتعدد للمتغيرات الثلاثة محل الدراسة يبلغ ٠,٩٣ لذلك فهو ارتباط قوى جدا وطردى.

ويفيد معامل الاتباط هذا فى ايجاد معامل التحديد وهو مربع معامل الارتباط أى هو $ر^{٢}$ وهو فى المثال يساوى ٠,٨٨.

ويفيد معامل التحديد فى تفسير هام وهو أن ٨٨% من التباين فى قيم ص (الاقتصاد) يمكن تفسيرة بالعلاقة الناتجة (المعادلة) بين الاقتصاد والاحصاء والرياضة ، وأن ١٢% من ذلك التباين يرجع الى الخطأ العشوائى ، لذلك يستخدم معامل التحديد هذا فى اجراء اختبار احسن المطابقة * لدالة الانحدار المقاسة ، أى يمكن استخدام هذا المعامل فى عملية الاختيار بين دوال الانحدار المختلفة للحالة موضع الدراسة .

وقد تميل صيغة معامل التحديد $ر^{٢}_{ص/س ى}$ الى تقدير معامل التحديد بأكبر من قيمته وذلك اذا كان عدد المشاهدات قليلة فى المثال الحالى أو يكون عدد المتغيرات فى دالة الانحدار كبير ، وهنا يجب تعديل معامل التحدد بالصيغة :

---

* test of goodness of fit .



$$\bar{ر}^{2}_{\text{ص/س ى}} = ١ - (ر^{2}_{\text{ص/س ى}}) \left(\frac{ن - ١}{ن - ك}\right)$$

حيث :-

ن : عدد المفردات المشاهدة

ك : عدد المتغيرات الكلية فى دالة الانحدار

وبتطبيق ذلك على المثال

$$\therefore \bar{ر}^{2}_{\text{ص/س ى}} = ١ - (٠,٨٨) \left(\frac{٥ - ١}{٥ - ٣}\right)$$

$$= ١ - ١,٧٦$$

$$= ٠,٧٦$$

وتسمى الصيغة $\bar{ر}^{2}_{\text{ص/س ى}}$ بمعامل التحديد المعدل

The adjust Coefficient of multiple of multiple determination

ثانيا : الارتباط الجزئى للعلاقة ذات الخط المستقيم:

باستعراض سريع لحالات الارتباط يتضح أنه إذا كانت الظاهرة موضوع الدراسة تتأثر بمتغيرين فقط فان قياس العلاقة بينهما تتم بمعامل الارتباط البسيط (الكلى)، أما إذا كانت الظاهرة تتأثر بأكثر من متغيرين أردنا قياس العلاقة بين تلك المتغيرات مجتمعة فانة يتم استخدام معامل الارتباط المتعدد ، وأما إذا كانت الظاهرة تتأثر بأكثر من متغيرين وأردنا قياس العلاقة بين متغيرين فقط مع ثبات المتغيرات الأخرى فانة يتم استخدام معامل الارتباط الجزئى .



واذا كنا بصدد ثلاث متغيرات ورمزنا الـــى المتغـــير الأول بـــالرقم ١ والمتغير الثانى بالرقم ٢ والمتغير الثالث بالرقم ٣ فان صيغـــة معـــامل الارتباط الجزئى هى :-

$$\text{ر}_{21.3} = \frac{\text{ر}_{21} - \text{ر}_{31}\,\text{ر}_{32}}{\sqrt{(1-\text{ر}^2_{31})(1-\text{ر}^2_{32})}}$$

حيث :

$\text{ر}_{21.3}$: معامل الارتباط البسيط بين المتغير رقم ١ والمتغير رقم ٢ مـــع ثبات المتغير رقم ٣ .

$$\text{ر}_{31.2} = \frac{\text{ر}_{31} - \text{ر}_{21}\,\text{ر}_{23}}{\sqrt{(1-\text{ر}^2_{21})(1-\text{ر}^2_{23})}}$$

$$\text{ر}_{32.1} = \frac{\text{ر}_{32} - \text{ر}_{12}\,\text{ر}_{13}}{\sqrt{(1-\text{ر}^2_{12})(1-\text{ر}^2_{13})}}$$

مثال

من بيانات المثال السابق على الارتباط المتعدد احسب معاملات الارتبـــاط الجزئى مع تفسير الناتج .

الحـــــل

كان المثال على ثلاث متغيرات وهى الاقتصاد والاحصاء والرياضـــــة ، وكانت رموزها ص ، س ، ى ، والآن رموزها ١ ، ٢ ، ٣ وعلية :

$$ر_{٢١.٣} = \frac{ر_{٢١} - ر_{٣١}\, ر_{٣٢}}{\sqrt{(١ - ر^{٢}_{٣١})(١ - ر^{٢}_{٣٢})}}$$

$$= \frac{٠,٩٢ - ٠,٧٨ \times ٠,٧٦}{\sqrt{(١ - ٠,٦١)(١ - ٠,٥٨)}}$$

$$= ٠,٨٢$$

أى أن العرقة بين الاقتصاد والرياضة والاحصـــــاء مـــع ثبـــات الرياضة علاقة قوية موجبة .

$$ر_{٣١.٢} = \frac{ر_{٣١} - ر_{٢١}\, ر_{٢٣}}{\sqrt{(١ - ر^{٢}_{٢١})(١ - ر^{٢}_{٢٣})}}$$

$$= \frac{٠,٧٨ - ٠,٩٢ \times ٠,٧٨٧}{\sqrt{(١ - ٠,٨٥)(١ - ٠,٨٥)}}$$

$$= ٠,٣٢$$

أى أن العلاقة بين الاقتصاد والرياضة مع ثبات الاحصاءعلاقة ضعيفة موجبة .



$$\text{ر}_{١.٣٢} = \frac{\text{ر}_{٣٢} - \text{ر}_{١٢}\ \text{ر}_{١٣}}{\sqrt{(١ - \text{ر}^{٢}_{١٢})\ (١ - \text{ر}^{٢}_{١٣})}}$$

$$= \frac{٠,٧٦ - ٠,٩٢ \times ٠,٧٨٧}{\sqrt{(١ - ٠,٨٥)\ (١ - ٠,٦١)}}$$

$$= ٠,١٨$$

أى أن العلاقة بين الاحصاء والرياضة مع ثبات الاقتصاد علاقة ضعيفة جداً موجبة .

ملحوظة :-

تبين أن معامل الارتباط الجزئى بين الاقتصاد والاحصاء مع ثبات الرياضة يساوى ٠,٨٢ وبين الإقتصاد والرياضة مع ثبات الاحصاء يساوى ٠,٣٢ ، وبين الاحصاء والرياضة يساوى ٠,١٨ ، إلا أن معامل الاتباط المتعدد أى للمتغيرات الثلاث معاً يساوى ٠,٩٣ مما يعنى أهمية المواد الثلاثة معاً وذلك حسب البيانات المعطاه .

الفصل الثانى

الانحدار المتعدد

الانحدار المتعدد للعلاقة ذات الخط المستقيم :

يستخدم الانحدار المتعدد من خلال تجميع البيانات المناسبة عن المتغيرات محل الدراسة فى التوصل إلى المعادلة الرياضية التى تحكم العلاقة بين أحد المتغيرات وهو المتغير التابع وباقى المتغيرات وهى المتغيرات المستقلة وذلك باستخدام طريقة المربعات الصغرى بهدف أما تفسير تلك العلاقة أو التنبؤ بسلوكها فى المستقبل . وقد أصبح واضحا أن الطريقة الجبرية هى الطريقة الوحيدة فى دراسة الانحدار المتعدد حيث استحالة العرض البيانى اللهم فى حالة ثلاث متغيرات فيكون العرض البيانى فى الفراغ ذو الثلاث ابعاد .

مثال /

اذا كنا بصدد دراسة العلاقة السببية بين الاقتصاد كمتغير تابع وكلا من الاحصاء والرياضة كمتغيرين مستقلين لطلبة السياحة والفنادق ، وتم تجميع بيانات عن درجات الامتحان فى الاقتصاد والاحصاء والرياضة لمجموعة مكونة من ٥ طلاب منهم فكانت البيانات كما فى بينات المثال السابق عن الارتباط المتعدد ، وبفرض أن هذه العلاقة ذات خط مستقيم فأوجد معادلة الانحدار المتعدد مع تفسير النتائج .



الحـل

$$ص = د (س ، ى)$$

حيث :

ص : هى ص١ ، ص٢ ، ص٣ ، .... درجات الاقتصاد

س : هى س١ ، س٢ ، س٣ ، ..... . درجات الاحصاء

ى : هى ى١ ، ى٢ ، ى٣ ، .... درجات الرياضة

، .. العلاقة هذه ذات خط مستقيم (فرضا)

.. ص = أ + ب س + جـ ى

، باستخدام البيانات المجمعة عن المتغيرات الثلاثة واستخدام طريقة المربعات الصغرى يتبين أن المعادلة الرياضية المقدرة التى تحكم هذه العلاقة هى* :

$$\hat{ص} = -٠,٨ + ١,٥٣ س + ٠,٦٧ ى$$

التفسير :

أن المتغير المستقل فى المعادلة وهو الاحصاء يؤثر على المتغير التابع ص وهو الاقتصاد تأثيرا موجبا لدرجة أن تغير الاحصاء بمقدار الوحدة يؤدى إلى تغير فى الاقتصاد قدر ١,٥٣ وحدة ، فى حين الأمر يختلف فى حالة المتغير ى ، ومن هذا يتضح أن الاحصاء أكثر تأثيرا على النجاح فى الاقتصاد .

---

* تم التوصل إلى هذه المعادلة عند دراسة الارتباط المتعدد .

مثال

الجدول التالى لبيانات تم تجميعها من مجموعة مكونة من ٦ سائحين عن الانفاق الكلى لكل منهم (ص) بالالف جنيه ، وإنفاق على الطعام (س) بالمائة جنيه ، وإنفاقه على الهدايا (ى) بالمائة جنيه :

ص	١٠	١٢	١٥	١٨	٢٠	٢٣
س	٢	٣	٥	٧	٨	١٠
ى	٣	٥	٨	١٠	١٢	١٣

والمطلوب :

- تقدير معادلة الانحدار التى تحكم هذه العلاقة بفرض أن هذه العلاقة ذات خط مستقيم مع تفسير نتائج المعادلة .
- باستخدام هذه المعادلة قدر الانفاق الكلى للسائح اذا بلغ انفاق السائح على الطعام والهدايا ١٥٠٠ جنيه ، ٢٠٠٠ جنيه على الترتيب .

الحل

.. ص = د (س ، ى)

، .. العلاقة ذات خط مستقيم (فرضا)

.. ص = أ + ب س + جـ ى

ولتوفيق أفضل خط انحدار ص /س ى ، يتم استخدام طريقة المربعات الصغرى أى المعادلات الطبيعية الثلاث التالية :

مج ص = ن أ + ب مج س + جـ مج ى

مجـ س ص = أ مجـ س + ب مجـ س² + جـ مجـ ى س

بالمثل

مجـ ى ص = أ مجـ س + ب مجـ س ى + جـ مجـ ى²

وهذا يستلزم تكوين الجدول الاحصائى التالى

س	ص	ى	ص س	ص ى	س ى	ص²	س²	ى²
١٠	٢	٣	٢٠	٣٠	٦	١٠٠	٤	٩
١٢	٣	٥	٣٦	٦٠	١٥	١٤٤	٩	٢٥
١٥	٥	٨	٧٥	١٢٠	٤٠	٢٢٥	٢٥	٦٤
١٨	٧	١٠	١٢٦	١٨٠	٧٠	٣٢٤	٤٩	١٠٠
٢٠	٨	١٢	١٦٠	٢٤٠	٩٦	٤٠٠	٦٤	١٤٤
٢٣	١٠	١٣	٢٣٠	٢٩٩	١٣٠	٥٢٩	١٠٠	١٦٩
٩٨	٣٥	٥١	٦٤٧	٩٢٩	٣٥٧	١٧٢٢	[illegible]	٥٠١

[illegible] جـ

٦٤٧ = ٣٥ أ + ٢١٥ ب + ٣٥٧ جـ

٩٢٩ = ٥١ أ + ٣٥٧ ب + ٥٩ جـ

وبحل هذه المعادلات الثلاث جبريا سواء بالحذف أو التعويض أو باستخدام المصفوفات فإن أ = ٦,٩٧ ، ب = ١,٦٥ ، جـ = −٠,٠٣

∴ ص = ٦,٩٧ + ١,٦٥ س − ٠,٠٣ ى

وهذه هى المعادلة الجبرية لأفضل خط مستقيم يمثل العلاقة بين بيانات المتغيرات الثلاثة موضوع الدراسة



وتعنى المعادلة أنه اذا زاد انفاق السائح على الطعام بمقدار الوحدة (مائة جنيه) زاد الانفاق الكلى للسائح بمقدار ١,٦٥ وحدة (ألف جنيه) أى لو زاد انفاق السائح على الطعام بمقدار واحد جنيه زاد الانفاق الكلى للسائح بمقدار ١٦,٥ جنيه ، فى حين لو زاد انفاق السائح على الهدايا بمقدار الوحدة (مائة جنيه) نقص الانفاق الكلى للسائح بمقدار ٠,٠٣ وحدة (ألف جنيه) أى لو زاد انفاق السائح على الهدايا بمقدار واحد جنيه نقص الانفاق الكلى للسائح بمقدار ٠,٣ جنيه . وهذا يعنى أن أى تحسن ولو طفيف فى مستوى المأكل المقدم للسائح سوف يؤدى إلى زيادة كبيرة فى الانفاق الكلى للسائح ، فى حين سيحدث العكس بالنسبة للهدايا المقدمة للسائحين .

- قيمة الانفاق الكلى للسائح عندما يكون انفاق السائح على الطعام ١٥ وحدة والهدايا ٢٠ وحده .

ص̂ = ٦,٩٧ + ١,٦٥ × ١٥ − ٠,٠٣ × ٢٠

= ٣١,١٢ وحدة

مثال

عند دراسة العلاقة بين ضغط الدم (ص) وكلا من العمر (س) والوزن (ى) لمجموعة مكونة من ٧ أشخاص متساوون فى الطول تقريبا حصلنا على البيانات التالية :

ص	١٢٠	١٢٤	١١٧	١٣٢	١٢٣	١٥٥	١٤٧
س	٥٠	٢٠	٣٠	٤٠	٥٥	٤٠	٢٠
ى	١٥٢	١٧١	١٥٨	١٧٠	١٥٣	١٩٠	١٨٥



والمطلوب :

- تقدير معادلة الانحدار التى تحكم هذه العلاقة بفرض أن هذه العلاقة ذات خط مستقيم مع تفسير نتائج المعادلة .
- باستخدام المعادلة الناتجة قدر قيمة ص لشخص وزنه ١٨٠ وعمره ٤٥ .

الحل

.. ص = د (س ، ى)

، .. العلاقة ذات خط مستقيم (فرضا)

.. ص = أ + ب س + جـ ى

ولتوفيق أفضل خط انحدار ص/س ى ، فإنه يتم استخدام طريقة المربعات الصغرى أى المعادلات الطبيعية الثلاث المعروفة :

تكوين الجدول الاحصائى اللازم :

س	ص	ى	ص س	ص ى	س ى	$ص^2$	$س^2$	$ى^2$
١٢,٠	٥,٠	١٥,٢	٦٠	١٨٢,٤	٧٦	١٤٤	٢٥	٢٣١,٠٤
١٢,٤	٥,٠	١٧,١	٢٤,٨	٢١٢,٠٤	٣٤,٢٠	١٥٣,٧٦	٤	٢٩٢,٤١
١١,٧	٣,٠	١٥,٨	٣٥,١	١٨٤,٨٦	٤٧,٤٠	١٣٦,٨٩	٩	٢٤٩,٦٤
١٣,٢	٤,٠	١٧,٠	٥٢,٨	٢٢٤,٤	٦٨	١٧٤,٢٤	١٦	٢٨٩
١٢,٣	٥.٥	١٥,٣	٦٧,٦٥	١٨٨,١٩	٨٤,١٥	١٥١,٢٩	٣٠,٢٥	٢٣٤,٠٩
١٥,٥	٤,٠	١٩,٠	٦٢	٢٩٤,٥	٧٦	٢٤٠,٢٥	١٦	٣٦١
١٤,٧	٢,٠	١٨,٥	٢٩,٤	٢٧١,٩٥	٣٧	٢١٦,٠٩	٤ -	٣٤٢,٢٥
١١٧,٩	٢٥,٢	٩١,٨	٣٣١,٧٥	١٥٥٨,٣٤	٤٢٢,٧٥	١٢١٦,٥٢	١٠٤,٢٥	١٩٩٩,٤٥

ملحوظة :-

لتجنب ظهور أرقام كبيرة فى العمليات الحسابية فقد تم قسمة بيانات كل من ص ، س ، ى على ١٠ بشرط أن يعالج ذلك فيما بعد كما يلى :

١٩,٨ = ٧ أ + ١١٧,٩ ب + ٢٥,٢ جـ

١٥٥٨,٣٤ = ١١٧,٩ أ + ١٩٩٩,٤٣ ب + ٤٢٢,٧٥ جـ

٣٣١,٧٥ = ٢٥,٢ أ + ٤٢٢,٧٥ ب + ١٠٤,٢٥ جـ

بحل هذه المعادلات الثلاث فإن أ = ٦,٨٧٠ ، ب = ٠,٤١٦ ، جـ = ١,٠٩٦

∴ ص = - ٦,٨٧ + ٠,٤١٦ س + ١,٠٩٦ ى .

وأنه عندما ى (الوزن) = ١٧٥ ، س (العمر) = ٤٤ فإنه يجب المعالجة لهذين البيانين أى ى = ١٧,٥ ، س = ٤,٤ وذلك قبل اجراء التعويض فى المعالدة .

∴ ص = - ٦,٨٧ + ٠,٤١٦ × ٤,٤ + ١,٠٩٦ × ١٧,٥ .

= ١٤,١٤

∴ ص للبيانات الأصلية = ١٤,١٤ × ١٠

= ١٤١,٤

وتفيد المعادلة أن الوزن يؤثر على ضغط الدم أكثر مما يؤثر العمر وذلك عند ثبات الطول .



الخطأ المعيارى للتقدير فى الانحدار المتعدد :-

يفيد الخطأ المعيارى للتقدير فى التعرف على مدى دقة التوفيق فى معادلة الانحدار التى تم التوصل إليها من البيانات المعطاه . وكما سبق إيضاحه بأنه كلما قل الخطأ المعيارى كلما زادت دقة التوفيق لمعادلة الانحدار الناتجة .

كما قد سبق الايضاح بأن الخطأ المعيارى للتقدير فى حالة متغيرين أحدهما تابع والآخر مستقل والذى يرمز له بالرمز

$$خ_{ص/س} = \sqrt{\frac{مج\ ص^{٢} - أ\ مج\ ص - ب\ مج\ س\ ص}{ن - ٢}}$$

، أما الخطأ المعيارى للتقدير فى حالة عدة متغيرات فيرمز له بالرمز $خ_{ص/س١\ س٢\ س٣}$ ويساوى .

$$= \sqrt{\frac{مج\ ص^{٢} - أ\ مج\ ص - ب\ مج\ ص\ س_{١}\ \ جـ\ مج\ ص\ س_{٢} - د\ مج\ ص\ س_{٣}}{ن - ك}}$$

وذلك اذا كنا بصدد دراسة تأثير ثلاث متغيرات مستقلة على المتغير التابع ، وحيث ن هى عدد المشاهدات ، ك هى عدد المتغيرات الداخلة فى النموذج ، وهى فى هذا النموذج تساوى ٤ أى ص ، $س_{١}$ ، $س_{٢}$ ، $س_{٣}$.

مثال

من معطيات المثال الموجود فى الانحدار المتعدد تبين أن معادلة الانحدار الناتجة كانت :



ص = - ٠,٨ + ١,٥٣ س + ٠,٦٧ ى

والمطلوب : ايجاد الخطأ المعيارى لتقدير ص^ بمعلوميته س ، ى

الحـــل

تكوين الجدول الاحصائى اللازم الأول :

ص^	- ٠,٨ + ١,٥٣ س + ٠,٦٧ ى	قيمة ص^
ص^١	- ٠,٨ + ١,٥٣ × ١ + ٠,٦٧ × ٢	١,٤٠
ص^٢	- ٠,٨ + ١,٥٣ × ١ + ٠,٦٧ × ٢	٣,٦٠
ص^٣		٤,٤٦
ص^٤		٦,٦٦
ص^٥		٨,٨٦
المجموع		٢٥

تكوين الجدول الاحصائى اللازم الثانى : -

ص	ص^	(ص-ص^)	(ص-ص^)٢	(ص-ص̄)	(ص-ص̄)٢	(ص^-ص̄)	(ص^-ص)٢	ص٢
٢	١,٤٠	٠,٦	٠,٣٦	-٣	٩	- ٣,٦	١٢,٩٦	٤
٣	٣,٦٠	- ٠,٦	٠,٣٦	-٢	٤	- ١,٤	١,٩٦	٩
٥	٤,٤٦	٠,٥٣	٠,٢٨	٠	٠	- .٣٦	.١٣	٢٥
٥	٦,٦٦	- ١,٥٦	٢,٧٢	٠	٠	١,٦٦	٢,٧٦	٢٥
١٠	٨,٨٦	١,١٢	١,٢٥	٥	٢٥	٣,٨٦	١٤,٩	١٠٠
٢٥		صفر	٤,٩٧	صفر	٣٨	.١٦	٣٢,٧٦	١٦٣

ونستنتج من الجدول السابق العلاقة التالية:

$$\text{مج}(\text{ص}-\bar{\text{ص}})^2 = \text{مج}(\text{ص}-\hat{\text{ص}})^2 + \text{مج}(\hat{\text{ص}}-\bar{\text{ص}})^2$$

$$38 = 5 + 33$$

ملاحظات على العلاقة السابقة:

١-يسمى $\text{مج}(\text{ص}-\bar{\text{ص}})^2$ بمجموع مربعات الاختلاف الكلى، وكثيرا ما يطلق علية بالتباين الكلى .

٢-يسمى $\text{مج}(\text{ص}-\hat{\text{ص}})^2$ بمجموع مربعات الاختلاف عن خط الانحدار، وكثيرا ما يطلق علية بالتباين غير المفسر.

٣-يسمى $\text{مج}(\hat{\text{ص}}-\bar{\text{ص}})^2$ بمجموع الباقى ، وكثيرا ما يطلق علية بالتباين المفسر.

وعليـــــــة

إذا أردنا إيجاد الخطأ المعيارى للتقدير في المعادلة المذكـورة فـى هـذا المثال نتبع الاتى:

$$\therefore \text{خ}_{\text{ص/س،ى}} = \sqrt{\frac{\text{مج}(\text{ص}-\hat{\text{ص}})^2}{\text{ن}-3}}$$

$$= \sqrt{\frac{5}{2}}$$

$$= 1{,}58$$

اذا اردنا الخطأ المعيار السابق بطريقة آخرى نتبع الآتى:



$$\therefore \text{خ}_{\text{ص/س ى}} = \sqrt{\frac{\text{مج ص}^{2} - \text{أ مج ص} - \text{ب مج ص س} - \text{جـ مجـ ص ى}}{\text{ن} - 3}}$$

$$= \sqrt{\frac{١٦٣ - \times - ٠,٨ \times ٢٥ - ١,٥٣ \times ٩٣ - ٠,٦٧ \times ٥٣}{٢}}$$

$$= \sqrt{\frac{٥}{٢}}$$

$= ١,٥٨$

ويلاحظ أنها طريقة أسهل من الطريقة السابقة.

العلاقة بين الخطأ المعيارى للتقدير ومعنوية النموذج ككل:

لدراسة هذه العلاقة يتم استخدام جدول تحيل التباين في الانحدار، وسوف يتم تناوله باختصار حيث يتم تناله بالتفصيل في الجزء الثاني من هذا الكتاب باذن الله.

جدول تحليل التباين **ANOVA**

مصدر التباين	درجات الحرارة	مجموع مربعات الانحرافات	متوسط مجموع مربعات الانحرافات	F	F.5
الانحدار	٢	٣٣	١٦,٥	٦,٣	١٩
الباقى	٢	٥	٢,٥		
الكلى	٤	٣٨			



ويتضح من تحليل الجدول السابق ان F المحسوبة أقل من F الجدولية عند مستوى معنوية ٠,٠٥ ، وعلى ذلك لايوجد انحدار معنوى بين المتغير التابع ص والمتغيرين المستقلين س ، ى .

مثال

الجدول التالى علاقة بين تغير تابع ص وعده متغيرات مستقلة $س_١$ ، $س_٢$ ، $س_٣$ ، $س_٤$

ص	$س_١$	$س_٢$	$س_٣$	$س_٤$
٣٦١	٠,٩٢	١٠,٢٩١	١,٠٤	٣٠,٦
٤٠٠	٠,٩٠	١٠,٤٦٣	١,٠٥	٣٢,٤
٤٣١	٠,٩٠	١٠,٦٧٤	١,٠٣	٣٤,١
٣٥٠	١,٠١	١٠,٩٧٢	١,٠	٢٩,٦
٦١٤	٠,٩٢	١١,٥٧٣	١,٢١	٣٩,٧
٤٥٣	٠,٩٨	١٢,٤١٧	٠,٩٧	٣٢,٠
٥٩٧	٠,٨٩	١٣,٠٠١	١,٢	٣٧,٨
٣١٢	١,٠٢	١٣,٥٦١	٠,٨٧	٢٩,١
٣٠٧	١,٠٩	١٣,٩٧٢	٠,٧٩	٢٦,٨
٧٩٧	٠,٨٩	١٤,٦٣٥	١,٣٢	٣٨,٢
٨٩٠	٠,٨٠	١٥,٢٣٩	١,٣٧	٤٠,٢
٩٢٧	٠,٧٩	١٦,٠٠٨	١,٤٣	٣٩,٦
٧٢٠	٠,٩٢	١٦,٥٤١	١,٠١	٣١,٧
٦٤٥	١,١٧	١٧,٤١٧	٠,٨٢	٣٢,١
٨٤١	٠,٨٢	١٧,٩٠٨	١,٠	٣١,٧
٩٢٥	٠,٨٥	١٨,٢٤٦	١,١٦	٣٢,٣



والمطلوب:

١-اوجد معادلة الانحدار على فرض أنها على الصوره:

ص = أ + ب س$_1$ + حـ س$_2$ + د س$_3$ + ع س$_4$

٢-اوجد الخطأ المعيارى للتقدير

٣-كون جدول تحليل التباين.

<u>ارشـــادات</u> :

١-المعادلة ستكون:

ص = -٦١٠,٤٠٣ - ٤١٥,١١٨ س$_1$ + ٠٥٧٨. س$_2$

+ ٢٩٠,٠٥٢س$_3$ + ١٤,١٠٧س$_4$

٢-الخطأ المعيارى خ$_{ص/س_1 س_2 س_3 س_4}$ = ٤٧,١٦١٥٢

أسئلة وتمارين

س١ : عرف ما يأتى بإيجاز :

* المفردة الاحصائية
* المتغير الاحصائى
* المتجمع الاحصائى الكبير
* المتجمع الاحصائى الصغير
* المتجمع الاحصائى المحدود
* المتجمع الاحصائى غير المحدود

س٢ : يفضل اسلوب العينة فى جمع البيانات الاحصائية عن اسلوب الحصر الشامل لماذا ؟

س٣ : اذكر الاخطاء الشائعة عند جمع البيانات الاحصائية .

س٤ : لماذا لا يفضل عرض البيانات الاحصائية فى صيغة كتابية ؟ وما هى الطرق المفضلة فى عرضها .

س٥ : الجدول الاحصائى هو ترتيب نظم للبيانات فى شكل صفوف وأعمدة بحيث يسهل قراءتها وفهم مضمونها ، توضح الشروط الواجب توافرها عند عمل جدول احصائى .

س٦ : يشترط فى العرض البيانى أن يتم تقسيم المحورين تقسيما مناسبا ، فما أهمية كلمة مناسب ؟

س٧ : الجدول التالى يبين حركة الركاب في مطار القاهرة الدولـــى فـــى عامى ١٩٩٥ ، ١٩٩٦ :

عد الركاب بالألف

حركة الركاب / السنة	قادمون	راحلون	عائدون	جمله
١٩٩٥	٢٦٠	٢٥٧	١٥٣	٦٧٠
١٩٩٦	٤٥٢	٤٥٤	٢٧٩	١١٨٥

والمطلوب : عرض هذه البيانات فى شكل بيانى مناسب .

س٨ : الجدول التالى يوضح بيانات عن التمويل العام لميزانيـــة احـــدى الدول و اوجه الاستخدام :

مصادر التمويل	ضرائب الرحل	ضرائب مهن حره	ضرائب مشروبات	ضرائب آخرى	ارباح احتكارية	دخول آخرى	الفائض أو العجز	الاجمالى
	%٢٣	%٢٥	%١٠	%١٧	%٣.٣	%١٨	%٣,٧+	%١٠٠
الاستخدامات	احتياطى عام	تمويل محلى	استصلاح أراضى	تطور صناعى	تعليم وثقافة	أمن قومى	سياحة وفنادق	
	%١٠	%١٧	%١٧	%١٠	%٢١,٢	%١٨	%٦,٨	%١٠٠

والمطلوب : عرض هذه البيانات بيانيا باستخدام الأعمدة المجزأه .

س٩ : متى نلجأ لعمل جدول توزيع تكرارى للبيانات الاحصائية ؟ صمـم صورة للجدول اللازم .

س١٠ : عند العرض الجدولى للبيانات الاحصائية فـــى جـــدول توزيـــع تكرارى يشترط أن يكون عدد الفئات مناسبا ، فمـــا أهميـــة كلمـــة مناسب .

س١١ : وضح كيف يمكن قياس مركز الفئة ؟ وما هو الافتراض الرياضى عند استخدام مركز الفئة فى العمليات الاحصائية ؟ وما مدى قناعتك فى هذا الافتراض ؟ ثم وضح كيف انتبه العالم شبرد لهذا الافتراض ؟

س١٢ : يوضح الجدول التالى درجات امتحان مادة الاحصاء لمجموعتين من طلبة السياحة والفنادق .

فئات الدرجات	التكرار المطلق		التكرار النسبى		التكرار المئوى	
	سياحة	فنادق	سياحة	فنادق	سياحة	فنادق
-٥٠	٣	٥				
-٥٣	٧	٦				
-٥٦	١٥	٧				
-٥٩	٢٠	١٠				
-٦٢	٤٢	١١				
-٦٥	٣٠	١٩				
-٦٨	١٧	١٧				
-٧١	٨	٨				
-٧٤	٨	٨				
المجموع	١٥٠	١٠٠				

والمطلوب :

١. استكمل البيانات الناقصة بالجدول .

٢. هل يتساوى التوزيع التكرارى المطلق لطلبة السياحة مع التوزيع التكرارى المطلق لطلبة الفنادق فى الفئات الثلاثة الأخيرة ؟ لماذا ؟

٣. كيف يمكنك المقارنة بين هذين التوزيعين التكراريين ؟

س١٣ : الجدول التالي تنقصه بعض البيانات انقله فى ورقة اجابتك :

الفئات	٣٠-	٣٨-	٤٦-	٥٤-	٦٢-	٧٠-	المجموع
التكرارات	٢٠	٣٢		٢٣		٦	١٤٣
تكرار متجمع صاعد	٢٠		٩٧		١٣٧	١٤٣	
تكرار متجمع صاعد مئوى							

والمطلوب :

١. أوجد البيانات الناقصة بالجدول .

٢. أعرض هذا الجدول بيانيا ودون ملاحظاتك .

س١٤ : البيانات الأتية توضح التقديرات التى حصل عليها أربعون طالبا فى امتحان مادة الاحاء :

جيد	جيد جدا	راسب	مقبول	جيد
راسب	مقبول	ممتاز	جيد	راسب
مقبول	جيد	مقبول	جيد جدا	مقبول
راسب	مقبول	راسب	مقبول	مقبول
مقبول	جيد جدا	مقبول	جيد	مقبول
ممتاز	مقبول	راسب	راسب	جيد
مقبول	راسب	جيد	مقبول	جيد جدا
راسب	جيد	ممتاز	جيد	راسب

والمطلوب :

وضع هذه التقديرات فى جدول تكرارى بسيط .

س١٥ : البيانات الآتية تمثل أوزان ٨٠ طالبا بإحدى الكليات بالكيلوجرام:

٨٠	٧٥	٦٢	٦٩	٥١	٦١	٧٢	٨٣
٦١	٦٢	٦٤	٧٣	٧٣	٦٨	٨٦	٨٤
٥٥	٧٠	٧٩	٧٨	٦٦	٨٩	٦٣	٦٥
٧١	٧٦	٦٦	٧٣	٥٧	٦٥	٦٥	٧١

٧٦	٧٧	٧٨	٨٥	٦٨	٦٨	٦٦	٥٩
٦٦	٧٠	٧٢	٧٢	٨٢	٦٣	٦٨	٨٦
٦٩	٨٤	٧٤	٦٤	٦٣	٦٩	٦٧	٦٥
٧٤	٦٢	٧٠	٧٩	٧٢	٧١	٥٦	٦٧
٧٧	٧٦	٧٥	٧٤	٧٨	٥٧	٦٧	٦٠
٨٢	٦٧	٦٤	٦١	٦٠	٦٩	٥٧	٨١

والمطلوب :

١. تصنيف هذه البيانات فى جدول تكرارى بسيط طبقا لفئات الوزن .

٢. رسم المدرج التكرارى ، المضلع التكرارى ، المنحنى التكرارى لهذا التوزيع .

٣. رسم المنحنى المتجمع الصاعد ومنه أوجد عدد الطلبة الذين يقل وزنهم عن ٦٢ كجم .

٤. رسم المنحنى المتجمع الهابط ومنه أوجد عدد الطلبة الذين بلغ وزنهم ٧٣ كجم فأكثر ثم أوجد نسبتهم الى جملة الطلاب .

س١٦ : الجدول التالى يوضح توزيع الايرادات الشهرية بالألف دولار لعينة من الفنادق فى إحدى محافظات ج.م.ع .

فئات الايرادات	٣٠-	٤٠-	٥٠-	٦٠-	٧٠-	٨٠-	٩٠-١٠٠	الجملة
عدد الفنادق	٤	١١	٢٠	٣٦	١٧	٨	٤	١٠٠

والمطلوب :

أولا : تمثيل هذه البيانات باستخدام :

- المدرج التكرارى .
- المضلع التكرارى .
- المنحنى التكرارى

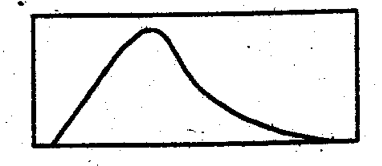
ثانيا : رسم المنحنى المتجمع الصاعد ومنه أوجد :

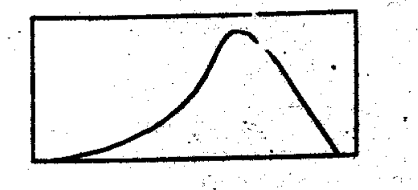
- عدد الفنادق التى تقل ايراداتها عن ٦٥ ألف دولار .
- الحد الأعلى للايرادات إلى حققها ٢٥ فندقا .

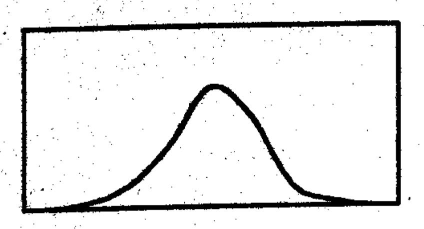
ثالثا : رسم المنحنى المتجمع النازل ومنه أوجد :

- عدد الفنادق التى بلغت فيها الايرادات ٧٠ ألف دولار .
- الحد الأدنى للإيرادات التى حققها ٦٠ فندقا .

س١٧ : اشرح توزيع حجم التكرارات على الفئات المناظرة لكل من التوزيعات التكرارية التالية ، ثم اذكر مثال من واقع الحياة لكل منها :







س١٨ : أوصف التوزيعات التكرارية فى الحالات التالية ووضحها بيانيا ، مع ذكر مثال من واقع الحياة لكل منها :

١. انخفاض التكرارات فى الفئات العليا .
٢. انخفاض التكرارات فى الفئات الدنيا
٣. ارتفاع التكرارات فى الفئات الدنيا .
٤. ارتفاع التكرارات فى الفئات العليا .
٥. تصاعد التكرارات ثم هبوطها بنفس المعدل على محور الفئات .

س١٩: تعلم أن مج س هو رمز للدلاله على مجموع قيم المتغير س حيث س هى : $س_{١}$ ، $س_{٢}$ ، $س_{٣}$ ، .... $س_{ن}$ فأجب عما يلى :

- مج أ س = ... حيث أ مقدار ثابت .
- مج أ = ... حيث أ مقدار ثابت لعدد من المرات .
- مج ($\frac{س_{ر}}{ص_{ر}}$) هل يساوى $\frac{مج\ س_{ر}}{مج\ ص_{ر}}$
- مج $س_{ر}$ $ص_{ر}$ هل يساوى (مج $س_{ر}$) (مج $ص_{ر}$)

س٢٠ : تعلم أن $\sum_{ر=١}^{ن}$ س ر هو أيضا رمز للدلالة على مجموع قيم المتغير $س_{ر}$ حيث ر هى ١ ، ٢ ، ٣ ، .... ن والمطلوب :

- $\sum_{ر=١}^{ن}$ أ $س_{ر}$ = ... حيث أ مقدار ثابت .
- $\sum_{ر=١}^{ن}$ أر = ... حيث أ مقدار ثابت .
- $\sum_{ر=١}^{ن}$ ٩ $س_{ر}$ $ص_{ر}$ = ...
- $\sum_{ر=١}^{ن}$ ($س_{ر}$ + ٧) = ...
- اثبت أن :

- مج ($أ_{١}$ س + $أ_{٢}$ ص $أ_{٣}$ ع) = $أ_{١}$ مج س + $أ_{٢}$ مج ص – $أ_{٣}$ مج ع

حيث : $أ_{١}$ ، $أ_{٢}$ ، $أ_{٣}$ مقادير ثابتة .

س٢١ : ما المقصود بالنزعة المركزية للبيانات الاحصائية .

س٢٢ : اذكر مقاييس النزعة المركزية للبيانات الاحصائية مبينا متى يفضل أستخدام كل منها عن الأخرى .

س٢٣ : أوجد كلما أمكن قيم المتوسط الحسابى والوسيط والمنوال والمتوسط الهندسي لكل من مجموعات البيانات التالية : -

١) ٢ ، ٥ ، ٧ ، ٣ ، ٤

٢) ٦ ، ٧ ، ٨ ، ٩ ، ٤ ، ٩

٣) ١٨ ، ٢٢ ، ٢٥ ، ٩٨ ، ١١ ، ١٥ ، ١١ ، ١٧ ، ١٠

٤) ١٠ ، -٨ ، ٩ ، ٧ ، ٥ ، ٦ ، ٤

٥) ١ ، صفر ، ٢ ، ٣ ، ٢

س٢٤ : كانت المبيعات لإحدى شركات صناعة أجهزة التليفزيون عام ٢٠٠٠ كالتالى :

التليفزيون بالبوصه	١٤	١٦	١٧	١٩	٢١	٢٣
المبيعات بالألف تليفزيون	١٦	٢٢	٢٩	٥١	١٤	٨

فما هو اتساع الشاشة بالبوصة الذى يمكنك أن تنصح الشركة بالاتجاه نحو إنتاجه خلال العام التالى .

س٢٥ : الجدول التالى يبين قيمة المبيعات اليومية لأحد المطاعم السياحية (بالجنيه)

الأيام	الأحد	الأثنين	الثلاثاء	الأربعاء	الخميس
قيمة المبيعات	٢٠٠٠	٢٥٠٠	٥٠٠٠	٧٥٠٠	١٠٥٠٠
نسبة الزيادة	-	١٢٥	٢٠٠	١٥٠	١٤٠

والمطلوب :

١. قدر متوس الزيادة فى حجـم المبيعـات باسـتخدام الوسـط الحسابى والوسط الهندسى ، وأيهما أفضل فى القياس ؟ولماذا؟
٢. بين باستخدام البيانات المعطاه أى الوسطين أفضل فى القياس.

س٢٦ : الجدول التالى يبين النقص فى قيمة أحد الأصول خلال الفـترة (٩٥-٢٠٠٠) .

السنة	١٩٩٥	١٩٩٦	١٩٩٧	١٩٩٨	١٩٩٩	٢٠٠٠
قيمة الأصل بالجنيه	٧٢٠	٦٠٩	٥٠٠	٣٧٥	٢٢٧	١٠٧

والمطلوب :-

إيجاد متوسط معدل النقص فى قيمة الأصل خلال هذه الفترة .

س٢٧ : أوجد الوسط الهندسى للتوزيع التكرارى المبين بالجدول الآتى:-

فئات	٣٥-	٤٥-	٥٥-	٦٥-	٧٥-	٨٥-	٩٥-	١٠٥-	١١٥-١٢٥	المجموع
تكرارات	١	٦	٨	٥	١٤	٢٢	٢٧	١٨	١٤	١١٥

س٢٨ : اذا كان عدد سكان إحدى مـدن ج.م.ع عـام ١٩٩٠ هـو ١,٦ مليون نسمة ، ثم فى عام ٢٠٠٠ بلغ ٣,٦ مليون نسـمة ، فقـدر عدد سكان هذه المدينة عام ١٩٩٥ .

س٢٩ : للبيانات التالية أحسب الوسط الحسابى والوسيط والمنـوال ثـم لاحظ مواقعهم على الرسم البيانى وقارن بينهم .

فئات	٢٥-	٣٥-	٤٥-	٥٥-	٦٥-	٧٥-٨٥	المجموع
تكرارات	٢	٤	٧	١٠	٩	٨	٤٠

س٣٠ : اثبت أن مجموع انحرافات القيم عن متوسطها الحسابى يساوى صفر ، ثم تحقق من صحة المتساوية : مجـ س – ن س = صفر.

س٣١ : اذكر أى مقاييس المتوسط أفضل فى قياس النزعة المركزية لمجموعة البيانات فى الحالات الآتية : –

(أ) إذا كان مجموعة البيانات قيمة شاذه .

(ب) إذا كان التوزيع التكرارى لمجموعة البيانات قريب من التوزيع المتماثل .

(جـ) إذا كان التوزيع التكرارى لمجموعة البيانات غير متماثل بدرجة كبيرة .

(د) إذا كانت الجداول التكرارية مفتوحة .

س٣٢ :اذكر المصطلح الأحصائى للعبارات التالية مع الايضاح البيانى:–

(أ) ميول البيانات للتجمع حول مركزها .

(ب) تباعد وتقارب البيانات .

(جـ) تجمع معظم البيانات عند الفئات العليا .

(د) تجمع معظم البيانات عند الفئات الدنيا

(هـ) تجمع نصف البيانات عند الفئات الوسطى وتجمع النصف الباقى عند الفئات الدنيا والعليا وذلك وفق متجمعات صاعده وهابطه متماثله .

(و) تجمع أكثر من نصف البيانات عند الفئات الوسطى وتجمع الباقى عند الفئات الدنيا والعليا وذلك وفق متجمعات صاعده وهابطه متماثله .

(ل) تجمع أقل من نصف البيانات عند الفئات الوسطى وتجمع الباقى عند الفئات الدنيا والعليا وذلك وفق متجمعات صاعده وهابطه متماثله .

س٣٣ : هل يكتمل الوصف الاحصائى للبيانات باستخدام مقاييس النزعة المركزية ؟ وضح ذلك بمثال من عندك .

س٣٤ : ما المقصود بتشتت البيانات الاحصائية ، واذكر فقط مقاييس التشتت .

س٣٥ : التوزيع التكرارى التالى هو توزيع سرعة الرياح بالعقدة فى ١٢٠ يوم بإحدى المدن :

سرعة الرياح	٢-	٨-	١٤-	٢٠-	٢٦-	٣٢-	٣٨-	٤٤-	المجموع
عدد الايام	٢	٨	١٨	٢٤	٣٠	١٨	١٢	٨	١٢٠

والمطلوب :-

احسب تشتت هذا التوزيع باستخدام مقاييس التشتت التالية :
المدى ، الانحراف الربيعى ، الانحراف المتوسط ، التباين ، الانحراف المعيارى .

س٣٦ : مجموع انحرافات القيم عن متوسطها الحسابى يساوى الصفر ، لكن فى التوزيعات التكرارية قد يكون مجموع انحرافات القيم عن متوسطها الحسابى لا يساوى الصفر لماذا ؟ .

س٣٧ : مجموع انحرافات القيم عن متوسطها الحسابى يساوى الصفر ، فماذا عن مجموع مربعات انحرافات القيم عن متوسطها الحسابى .

س٣٨ : ماذا عن قيمة كلا من :

(أ) متوسط مجموع انحرافات القيم عن متوسطها الحسابى .

(ب) متوسط مجموع مربعات القيم عن متوسطها الحسابى .

س٣٩ : اثبت أن :

$$\frac{\text{مج}\,(\text{س} - \bar{\text{س}})^2}{\text{ن}} = \frac{\text{مج}\,\text{س}^2 - \frac{(\text{مج}\,\text{س})^2}{\text{ن}}}{\text{ن}}$$

س٤٠ : طلب منك الأتى :

(أ) المقارنة بين تشتت مجموعتين مختلفتين فى وحدات القياس

(ب) مقارنة مفردة تنتمى إلى مجموعتين مختلفتين .

فماذا أنت فاعـــل .

س٤١ : بدراسة توزيعين تكرارين عن ظاهرتين مختلفتين تبين الأتى :-

الظاهرة الأولى : وسطها الحسابى = ٧٥ ،

، انحرافها المعيارى = ١٥

الظاهرة الثانية : بياناتها كالتالى :

فئات	١٨−	٢٨−	٣٨−	٤٨−	٥٨−٦٨
تكرارات	٢٦	٥٩	١٠٥	٧٩	٣١

فأى الظاهرتين أكثر تشتتا ؟

س٤٢ : فى عام ٢٠٠٠ كان متوسط المبيعات لإحدى شركات الســـياحة هو ١٥٢ ألف دولار بانحراف معيارى ١١ ألـــف دولار ، بينمـــا متوسط المبيعات لشركة سياحية أخرى كـــان ٢٠٠ ألــف دولار

بانحراف معيارى ١٨ ألف دولار ، فقارن مدى التارجح فى مبيعات الشركتين .

س٤٣ : التالى هو العرض البيانى لتوزيع درجات امتحان مادتى الاحصاء والاقتصاد لعدد ١٢٠ طالب :

والمطلوب :

(أ) ما حكمك على اداء الطالب فى الامتحانين لو استندت فقط فى المقارنة على المتوسط الحسابى لكلا الامتحانين ، وهل حكمك هذا يتفق مع ما يوضحه التوزيع البيانى للمتجمعين .

(ب) اذا علمت أن الانحراف المعيارى للاحصاء والاقتصاد هو ١٥ ، ٩ على الترتيب ، وأن أحد الطلاب حصل على ٦ درجات فى المادتين فهل يمكنك القول أن مستوى الطالب فى الامتحانين واحد ؟ ولماذا ؟

(جـ) أوصف احصائيا مستوى الاسئلة فى امتحان الاحصاء والاقتصاد .

س٤٤ : حول القيم الأتية إلى درجات معيارية :

٦ ، ٢ ، ٨ ، ٧ ، ٥

س ٤٥ : اذا كان العزم الأول حول الوسط الحسابى يساوى الصفر ، فماذا عن العزم الثانى حول الوسط الحسابى ؟

س٤٦ : اذا كان العزم الثانى حول الوسط الحسابى لمجموعة البيانات يساوى ٤٩ فما قيمة التباين والإنحراف المعيارى لهذه البيانات .

س٤٧ : أوجد العزم الأول والثانى والثالث والرابع لمجموعة الارقام التالية :

٢ ، ٣ ، ٧ ، ٨ ، ١٠

س٤٨ : أوجد العزم اللأول والثانى والثالث و لرابع حول النقطة ٤ لمجموعة البيانات فى السؤال ٤٧ .

س٤٩ : أوجد العزم الأول والثانى والثالث والرابع حول المتوسط الحسابى لمجموعة البيانات فى السؤال ٤٧ ، وماذا تلاحظ .

س٥٠ : الجدول التالى هو توزيع تكرارى لأوزان ١٠٠ طالب بإحدى الكليات :

فئات	٦٠-٦٢	٦٣-٦٥	٦٦-٦٨	٦٩-٧١	٧٢-٧٤	المجموع
تكرارات	٥	١٨	٤٢	٢٧	٨	١٠٠

والمطلوب :

(أ) أوجد العزم الأول والثانى والثالث والرابع حول الوسط الحسابى لهذا التوزيع .

(ب) أوجد معامل الالتواء وبين نوع التوزيع من حيث الالتواء .

(جـ) بين درجة تفرطح التوزيع .

(د) استخدام نتائج أ ، ب ، جـ في وصف هذا التوزيع احصائيا

س٥١ : توافرت لديك المعلومات التالية عن أحدى التوزيعات :

الانحراف المعيارى = ٥,٣٥٠ ، العزم الثالث=١٠١,٣٨٤

، العزم الرابع = ٢٥١٢,٩١٣

والمطلوب :

(أ) أوجد معامل الالتواء وبين نوعه .

(ب) أوجد معامل التفرطح وقارن بالتوزيع المعتدل المعيارى .

س٥٢ : ابت أن :

$$\frac{\text{مج}(\text{س} - \bar{\text{س}})^{2}\,\text{ك}}{\text{مج ك}} = \frac{\text{مج ح}^{2}\,\text{ك}}{\text{مج ك}} - \left(\frac{\text{مج ح ك}}{\text{مج ك}}\right)^{2}$$

س٥٣ : ماذا تقول فى التوزيع التكرارى إذا كان :

أولا : (أ) معامل التفرطح يساوى ٣

(ب) أو معامل التفرطح أقل من ٣

(جـ) أو معامل التفرطح أكبر من ٣ .

ثم وضح ذلك بيانيا مع الوصف الاحصائى .

ثانيا : (أ) معامل الالتواء يساوى صفر

(ب) أو معامل الالتواء يساوى -٠,٥

(جـ) أو معامل الالتواء يساوى ٠,٣

ثم وضح ذلك بيانيا مع الوصف الاحصائى .

س٥٤ : يكتمل الصف الاحصائى للتوزيع التكرارى بقياس كل من المتوسط ، التشتت ، الالتواء ، التفرطح ، فما هو وصفك الاحصائى للتوزيع المعتدل وأيضا التوزيع المعتدل المعيارى .

س٥٥ : التالى هو جدول توزيع تكرارى ما :

فئات	٤٠-	٤٢-	٤٤-	٤٦-	٤٨-	٥٠-	٥٢-	٥٤-	٥٦-	المجموع
تكرارات	٢١	٣٥	٥٢	٧٣	٩١	٨٤	٦٩	٤٥	٣٠	٥٠٠

والمطلوب :

(١)ارسم المنحنى التكرارى لهذا التوزيع مُهدا بقدر الامكان .

(٢)أوجد المنوال من الشكل البيانى .

(٣)ارسم على نفس الشكل المنحنى المتجمع الصاعد ومنه أوجد الوسيط والربيعين .

(٤)أوجد الانحراف المعيارى .

(٥)أوجد درجة الالتواء وبين نوعه لهذا التوزيع .

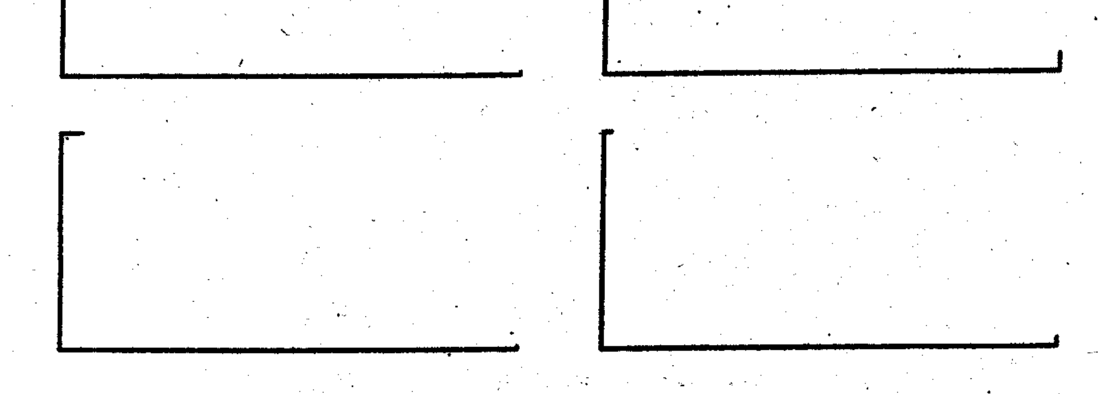
(٦)احسب العزمين الثالث والرابع لهذا التوزيع واستخدمها فى إيجاد الالتواء والتفرطح ثم الوصف الاحصائى للتوزيع .

س٥٦ : عند دراسة العلاقة بين متغيرين بيانيا فإن الأمر يتطلب عمل شكل الانتشار وخط الإنتشار فما المقصود بعمل كلا منهما ؟

س٥٧ : يتوقف على ميل خط الانتشار نوع الارتباط بين المتغيرين ، وضح ذلك ؟

س٥٨ : يتوقف على تقارب أو تباعد النقط حول خط الانتشار درجة الارتباط بين المتغيرين ، بين ذلك .

س٥٩ : الاشكال البيانية التالية هى لعلاقة بين متغيرين فبين نوع ودرجة الالتواء بمجرد النظر .



س٦٠ : عند قياس الارتباط البسيط جبريا يتم استخدام الانحراف المعيارى لماذا .

س٦١ : اكتب قانون معامل الارتباط بصيغتين .

س٦٢ : أثبت أن :

$$\frac{\text{مج}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}})}{\text{ن}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}}} = \frac{\text{مج س}^{٢} - \frac{\text{مج س} .\ \text{مج ص}}{\text{ن}}}{\sqrt{\left(\text{مجـ س}^{٢} - \frac{\text{مج}(\text{س})^{٢}}{\text{ن}}\right)\left(\text{مجـ ص}^{٢} - \frac{\text{مج}(\text{ص})^{٢}}{\text{ن}}\right)}}$$

س٦٣ : اذا كان الارتباط البسيط هو علاقة بين متغيرين فماذا عن الانحدار المتعدد والانحراف الجزئى .

س٦٤ : عند قياس الارتباط بين متغيرين يتم اختيار معامل الارتباط المناسب حتى يكون القياس سلبى فوضح ما هو معامل الارتباط المناسب فى الحالات التالية :

(أ) اذا كان خط الانتشار مستقيم

(ب) اذا كان خط الانتشار منحنى

(جـ) اذا كان خط الانتشار منحنى يخضع لمعادلة قوى أو معادلة أسية .

س٦٥ : انقل الجدول التالى فى ورقة إجابتك :

س	ص	(س-س)	(ص-ص)	(س-س)^2	(ص-ص)^2	(س-س)(ص-ص)
١	١					
٣	٢					
٤	٤					
٦	٤					
٨	٤					
٩	٧					
١١	٨					
١٤	٩					

المطلوب :-

(أ) استكمل البيانات الناقصة بالجدول .

(ب) أوجد الانحراف المعيارى لكلا من س ، ص وكذا التباين لكلا منهما .

(جـ) أوجد التغاير بين س ، ص .

(د) أوجه الارتباط الخطى بين س ، ص مبينا نوعه ودرجته .

س٦٦ : الجدول التالى يوضح التوزيع التكرارى لدرجات ١٠٠ طالب في مادتى الرياضة والاحصاء :

الرياضه / الاحصاء	(٤٠-٤٩)	(٥٠-٥٩)	(٦٠-٦٩)	(٧٠-٧٩)	(٨٠-٨٩)	(٩٠-٩٩)	التوزيع التكرارى للاحصاء
٤٠-٤٩				٢	٤	٤	١٠
٥٠-٥٩			١	٤	٦	٥	١٦
٦٠-٦٩			٥	١٠	٨	١	٢٤
٧٠-٧٩	١	٤	٩	٥	٢		٢١
٨٠-٨٩	٣	٦	٦	٢			١٧
٩٠-٩٩	٣	٥	٤				١٢
التوزيع التكرار للرياضه	٧	١٥	٢٥	٢٣	٢٠	١٠	١٠٠

والمطلوب :-

أوجد معامل الارتباط بين متغيري الرياضة والاحصاء لمجموعة الطلاب محل الدراسة .

س٦٧ : فيما يلى تقديرات عشرة طلاب فى مادتى الاقتصاد والاحصاء والمطلوب أحسب معامل الارتباط المناسب بين تقديرات المادتين :

تقديرات الاحصاء	جيد	ممتاز	جيد	مقبول	ضعيف	جيد	ضعيف جدا	ضعيف	جيد جدا	جيد جدا
تقديرات الاقتصاد	جيد	مقبول	جيد جدا	جيد	ضعيف جدا	جيد جدا	ضعيف	مقبول	ممتاز	جيد جدا

س٦٨ : لدراسة العلاقة بين الكمية المطلوبة ص بالكيلوجرام والسعر (س) بالجنيه كانت لدينا النتائج التالية :

مجـ س = ٦٠ مجـ ص = ٧٠ مجـ س ص = ٣٧٤

مجـ س^{٢} =٤٠٦ مجـ ص^{٢} = ٥٣٦ ن = ١٠

والمطلوب : احسب معامل الارتباط بين الكمية المطلوبة وسعرها .

س٧٠ : الجدول التالى يوضح بيانات عن الحالة الصحية للطالب ودخل اسرته .

الدخل / الحالة الصحية	(-٢٥٠)	(٢٥٠-)	(٥٠٠-)	(١٠٠٠-)	التوزيع التكرارى للحالى الصحية
جيده	-	٤	٦	٢	١٢
فوق متوسط	١	٦	١٤	٩	٣٠
متوسطه	٤	١٠	٨	٨	٣٠
دون المتوسط	٥	١	٢	-	٨
التوزيع التكرارى للدخل	١٠	٢١	٣٠	١٩	٨٠

والمطلوب :

قياس الارتباط بين المتغيرين

س٧١ : الجدول التالى يبين الايرادات السياسية (ص) بالمليون جنيه ، والخدمات الفندقية (س) بالألف جنيه ، ومصاريف الدعاية (ى) بالالف جنيه :

ص	١٠	١٢	١٥	١٨	٢٠	٢٣
س	٢	٣	٥	٧	٨	١٠
ى	٣	٥	٨	١٠	١٢	١٣

والمطلوب :

(أ) احسب معامل الارتباط المتعدد بغرض أن العلاقة بين المتغيرات الثلاثة خطيه مع تفسيرك للنتائج .

(ب) احسب معاملات الارتباط الجزئية مع تفسيرك للنتائج .

س٧٢ : البيانات التالية تمثل الاستثمارات س١ ، الدخول س٢ ، والمدخرات س٣ لعينة عشوائية مكونة من ٥ فنادق .

س١	١٢٠٠	٦٠٠٠	٦٠٠٠	٣٠٠٠	١٨٠٠٠
س٢	٨٠٠٠	١١٠٠	٩٠٠٠	٦٠٠٠	٦٠٠٠
س٣	٦٠٠	١٢٠٠	١٠٠٠	٧٠٠	٣٠٠

والمطلوب :-

(أ) احسب الارتباط الخطى المتعدد ل١ . ل٢ . ل٣ .

(ب) احسب معامل التحديد المتعدد ، ومنه أوجد نسبة التباين المفسر والتباين غير المفسر فى س١ على أساس العلاقة الخطية بين س١ ، س٢ ، س٣ .

س٧٣ : اذا كان الارتباط البسيط هو علاقة بين متغيرين وأيضا الانحدار البسيط هو علاقة بين متغيرين فما الفرق بينهما إذا ؟

س٧٤ : ماذا يعنى توفيق خط انحدار يمثل البيانات التى تم تجميعها عن المتغيرين محل الدراسة ؟

س٧٥ : التالية هى اشكال الانتشار للعلاقة الانحدارية ص/س للمتغيرين س ، ص :

والمطلوب :-

(أ) وفق خط انحدار مناسب للحالات الثلاثة بمجرد النظر ، مع كتابة المعادلة الرياضية المناسبة لكل منها ؟

(ب) ما مدى قناعتك فى التحليل البعدى إذا ما اعتمدت على خط الانحدار فى الخطة (أ) .

س٧٦ : اكتب المعادلات الطبيعية إذا كانت المعادلة الرياضية المناسبة لخط الانحدار هى :

(أ) ص = أ + ب س

(ب) ص = أ + ب س + جـ $س^{٢}$

(جـ) ص = أ + ب س + جـ $س^{٢}$ + د $س^{٣}$

(د) ص = أ $ب^{س}$

س٧٧ : ارسم شكل الانتشار لمجموعة البيانات :

س	٢	٤	١	٣	٥	٦
ص	٢	٥	٦	٣	٤	٢

ثم أوجد :

(أ) خط انحدار ص/س ومنه قدر ص عندما س = ٣,٥

(ب) خط انحدار س/ص ومنه قدر قيمة س عندما ص = ٢,٣

(جـ) اثبت أن نقطة تقاطع خط انحدار ص/س مع خط انحدار س/ص السابقين هى (س ، ص) لهما وذلك جبريا وبيانيا .

س٧٨ : إذا كانت العلاقة بين المتغيرين س ، ص تحتلها المعادلة ص = أ + ب س فأوجد خط انحدار ص/س إذا علمت أن :

الانحراف المعيارى لـ ص ($ع_{ص}$) = ١٥

: ، الانحراف المعياري لــ س ($ع_س$) = ٧

، معامــــل الارتبـــاط (ر) = ٠,٨

، المقـــــدار الثـــابت = ٢١

س٧٩ : الجدول التالي يوضح السن س وضغط الدم ص لعدد ١٢ امرأة :

س	٥٦	٤٢	٧٢	٣٦	٦٣	٤٧	٥٥	٤٩	٣٨	٤٢	٦٨	٦٠
ص	١٤٧	١٢٥	١٦٠	١١٨	١٤٩	١٢٨	١٥٠	١٤٥	١١٥	١٤٠	١٥٢	١٥٥

والمطلوب :-

(أ) أوجد معامل الارتباط بين ص ، س

(ب) أوجد معادلة إنحدار ص/س باستخدام المربعات الصغرى .

(جـ) قدر ضغط الدم لأمرأة عمرها ٤٥ سنة .

س٨٠ : باستخدام البيانات التالية مطلوب إيجاد معادلة انحدار ص/س .

س	ص	$س^٢$	س ص
-٢	٤,٧	٤	- ٩,٤
-١	٥,٢	١	- ٥,٢
صفر	٥,٨	صفر	صفر
١	٦,١	١	٦,١
٢	٦,٥	٤	١٣,٠٠

س٨١ : باستخدام المعادلة الناتجة من المسألة ٨٠ استكمل الجدول التالى:

ص المقدرة ($\hat{ص}$)	المعادلة الناتجة	قيمة ($\hat{ص}$)
$\hat{ص}_١$		
$\hat{ص}_٢$		
$\hat{ص}_٣$		
$\hat{ص}_٤$		
$\hat{ص}_٥$		

س٨٢ : من السؤال ٨٠ ، ٨١ اكمل الجدول التالى :

ص	$\hat{ص}$	(ص - $\hat{ص}$)	$(ص - \hat{ص})^٢$

ثم أجب عما يلى :

(أ) ماذا ترى فى مجـ ص ، مجـ $\hat{ص}$

(ب) ما قيمة مجـ (ص - $\hat{ص}$) ، وماذا ترى فى ذلك من خواص احصائية

(جـ) ما قيمة $\frac{مجـ (ص - \hat{ص})^٢}{ن - ٢}$ ، وماذا ترى فى ذلك من خواص احصائية

(د) ما قيمة $\sqrt{\frac{مجـ (ص - \hat{ص})^٢}{ن - ٢}}$ ، وما اسمه الاحصائى .

س٨٣ : اثبت أن :

$$\sqrt{\frac{\text{مج}\,(\text{ص}-\hat{\text{ص}})^2}{\text{ن}-2}} = \sqrt{\frac{\text{مج ص}^2 - \text{أ مج ص} - \text{ب مج س ص}}{\text{ن}-2}}$$

ثم تحقق من صحة هذه المتساوية فى المسألة ٨٠ .

س٨٤ : قارن بين قيمة (د) فى المسألة ٨٢ وقيمة نصف معامل خط انحدار ص/س الناتج فى المسألة ٨٠ موضحا رأيك فى معنوية هذه العلاقة .

س٨٥ : أوجد الخطأ المعيارى لعامل خط انحدار ص/س فى المسألة ٨٠

س٨٦ : اختبر معنوية معامل انحدار ص/س الناتج فى المسألة ٨٠ .

س٨٧ : فى ضوء ما توفر لديك من معلومات كاملة عن خط انحدار ص/س الناتج فى المسألة ٨٠ فما مدى قناعتك فى دقة توفيق خط انحدار ص/س الناتج فى المسألة ٨٠ ، وهل بعد ذلك يمكنك استخدام هذه المعادلة فى التفسير أو التنبؤ .

س٨٨ : ارسم شكل الانتشار لمجموعة البيانات :

س	٢٠	٣٠	٤٠	٥	٦٠	٧٠
ص	٥٤	٩٠	١٣٨	٢٠٦	٢٩٢	٣٩٦

ثم أوجد :

(أ) أفضل قطع مكافى على الصورة ص= أ + ب س + جـ س^2

(ب) قدر قيمة ص عندما س = ٨٠ ، عندما س = ٤٥

ملحوظه :

لتجنب ظهور أرقام كبيرة أثناء تطبيق طريقة المربعات الصغرى فإنه يمكنك قسمة بيانات س ، ص على الرقم ١٠ على أن يراعى ذلك أثناء التقدير .

س٨٩ : الجدول التالى علاقة بين متغيرين س/ص

س	٠,٩	١,٢	٢,٩	٣,١	٣,٩	٦,٢	١٢,٦	٢٥,٨
ص	١٨	٢٦	٣٢	٣٦	٣٥,٦	٥٧,٥	٨٤,٠	٨٧,٥

والمطلوب :-

(أ) ارسم شكل الانتشار مرة آخرى باستخدام التحويل لوص ، لـــوس ، وهل يمكن تقريب شكل الانتشار الجديد بخط مستقيم .

(ب) باستخدام التحويل السابق أوجد افضل خط مستقيم على الصورة :

لوص = أ + ب لوس

(د) اكتب العلاقة التى تربط المتغيرين س ، ص .

إرشادات فى الحل :

يلاحظ أن الخط المناسب للبيانات الأصلية يأخذ المعادلة ص= أ س^{ب} ، ولتقدير أفضل خط يمثل البيانات فى هذه الحالة يتم استخدام الصورة لوص = أ + ب لوس وعلى ذلك فإن الأمر يستلزم تكوين الجدول الاحصائى اللازم للحصول على قيم :

مجـ لوس ، مجـ لوص ، مجـ لوس لوص ، مجـ لوس^{٢}

، ثم التعويض فى المعادلتين الطبيعيتين :

مجـ لوص = ن أ + ب مجـ لوس (١)

مجـ لو س لوص = أ مجـ لوس + ب مجـ لو س^{٢} (٢)

وعلى ذلك ستكون المعادلة المقدرة هى لوص = ١,٣٢٣ + ٠,٤٧٢ لـوس ، ومن ثم فالمعادلة المقررة للبيانات الأصلية ستكون ص = $٢١ س^{٠,٤٧٢}$

س٩٠ : الجدول التالى يبين تطور عدد السكان فى إحدى المدن :

السنوات	١٩٩٠	١٩٩١	١٩٩٢	١٩٩٣	١٩٩٤	١٩٩٥	١٩٩٦	١٩٩٧	١٩٩٨	١٩٩٩	٢٠٠٠
عدد السكان	٠,٧٣١	٢,٣	٢,٦٣٧	١٦,١٥٩	١٧,٧	٣٩,٥٧٨	٧٤,٣٦١	١٤٧,٩٩٥	٢٠٣,٣٤١	٣٣٤,٣٨٧	٥٧٣,٢٢٤

والمطلوب :-

(أ) ارسم شكل الانتشار ، ثم ماذا تلاحظ .

(ب) ما رأيك فى شكل الانتشار والمعادلة ص = $أ ب^{س}$

(جـ) قدر خط انحدار ص/س ، ثم تنبأ بعدد السكان عام ٢٠١٠ .

إرشادات فى الحل :

استخدم العلاقة نصف اللوغاريتمية أى لوص = أ +ب س ، ثـم استخدم المعادلتين الطبيعيتين

مجـ لوص = ن أ + ب مجـ س

مجـ لوص = أ مجـ س + ب مجـ $س^{٢}$

وعلى ذلك ستكون المعادلة المقدرة لخط انحدار ص/س هى :

لوص = ٣,٠٦٠٦ + ٠,٢٨٥ س

ومن ثم فمعادلة خط انحدار ص/س للبيانات الصلية ستكون

ص = ١١٥٠ $(١,٩٢٧)^{س}$

س٩١ : وفق أفضل معادلة لمجموعة البيانات :

ص	٢	٥	٧	٨	٥
س	٨	٨	٦	٥	٣
ع	٠	١	١	٣	٤

وذلك إذا كانت العلاقة تأخذ الدالة ص= (س ، ع) وتأخذ المعادلة

ص = أ + ب س + ط ع

<u>إرشادات فى الحل :</u>

استخدم المعادلات الطبيعية التالية :-

مجـ ص = ن أ + ب مجـ س + جـ مجـ ع

مجـ س ص = أ مجـ س + ب مجـ $س^{٢}$ + جـ مجـ س ع

مجـ ص ع = أ مجـ ع + ب مجـ س ع + جـ مجـ $ع^{٢}$

س٩٢ : البيانات التالية عن الايرادات السياحية والخدمات الفندقية ومصاريف الدعاية :

الايرادات السياحية بالمليون	٢	٣	٥	٥	١٠
تكاليف الاقامة بالألف	١	٢	٣	٤	٥
مصاريف الدعاية بالألف	١	٢	١	٢	٣

والمطلوب :

تقدير معادلة خط الانحدار موضحا اثر تكاليف الاقامة ومصاريف الدعاية على الايرادات السياحية وذلك بفرض أن العلاقة معنوية .

س٩٣ : فى دراسة للعلاقة بين الإنفاق الكلى للسائح وإنفاقه على الطعام وعلى الهدايا كانت البيانات التالية :-

الانفاق الكلى للسائح	١٠	١٢	١٥	١٨	٢٠	٢٣
إنفاقه على الطعام	٢	٣	٥	٧	٨	١٠
انفاقه على الهدايا	٣	٥	٨	١٠	١٢	١٣

والمطلوب :-

(أ) تقدير معادلة خط الانحدار .

(ب) احسب الخطأ المعيارى فى التقدير .

(جـ) اختبر معنوية الانحدار للنموذج ككل .

إرشادات فى الحل :

معادلة خط الانحدار هى ص^ = ٠,٨٠ + ١,٥٣ س + ٠,٦ع ، الخطأ المعيارى فى التقدير هو ١,٦١ ، ولاختبار معنوية الانحدار للنموذج ككل يستخدم جدول تحليل التباين والذى سيرد شرحه بالتفصيل فى الجزء الثانى لهذا لكتاب بإذن الله .

س٩٤ : اذا كانت العلاقة بين المتغيرين س ، ص تحكمها القيم التالية :

ر = ٠,٨ ، عس = ٣ ، عص = ٢,٥ ، س= ١٥ ، ص = ٢٠

فالمطلوب :

(أ) أوجد معادلة خط انحدار ص/س ، معادلة خط انحدار س/ص .

(ب) احسب معامل الارتباط

إرشادات فى الحل :

ص^ = ٥+س ، س^ = ٢.٢ + ٠,٦٤ ص ، ر = √(١×٠,٦٤)

س٩٥ : اذا كانت ر = −٠,٧٥ ، س = ٥٠ ، ص = ٧٠٠
، عس = ١٤ ، عص = ١٠

المطلوب :-

(أ) احسب قيمة ص^ عندما س = ٦٥

(ب) احسب قيمة س^ عندما ص = ٤٠٠

س٩٦ : اذا كـــــانت ص^ = ٢٠ + ٠,٤ س ، س^ = ٦,٥ + ١.٤ ص
فاحسب معامل الارتباط .

س٩٧ : اذا كانت ر = ± ١ فأوجد قيمة خ٢ ص/س أى الخطــأ المعيــارى لتقدير ص بمعلومية س .

س٩٨ : اذا كانت العلاقة الانحدارية هى ص = د(س) وأخـــذ المتغــير المستقل عامل الزمن ، فماذا تسمى هذه العلاقة .

س٩٩ : يعطى الجدول الآتى قيمة المبيعات السنوية لإحــدى الشــركات السياحية

السنوات	١٩٩٢	٩٣	٩٤	٩٥	٩٦	٩٧	٩٨	٩٩	٢٠٠٠	٢٠٠١	٢٠٠٢
القيمة بالألف جنيه	٤	٧	٧	٨	٩	١١	١٣	١٤	١٧	١٩	٢٠

والمطلوب :

قدر خط الاتجاه العام على الصورة ص = أ ب س

س١٠٠:يوضح الجدول التالى تكاليف الإنتاج بقسم [illegible] لأحد الفنادق:

السنوات	١٩٩٧	٩٨	٩٩	[illegible]	٢٠٠١	٢٠٠٢
التكاليف بالمليون جنيه	٠,٩	١,٠	١,٣	[illegible]	[illegible]	٥,٠

والمطلوب :

(أ) تقدير خط العام على الصورة ص = أ + ب س + جـ $س^{2}$

(ب) اختبر دقة توفيق الخط البيانى الممثل للبيانات .

(جـ) تنبأ بقيمة التكاليف عام ٢٠٠٧ .

(د) هل يمكنك التنبؤ بقيمة التكاليف عام ٢٠١٠ ؟ ولماذا ؟